



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 126551

TO: Phuong Bui
Location: REM-2A15/3C18
Art Unit: 1638
Friday, July 09, 2004

Case Serial Number: 09/857525

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Bui,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 01:29:12 ; Search time 6278 Seconds
(without alignments)
10880.634 Million cell updates/sec

Title: US-09-857-525C-1

Perfect score: 1576

Sequence: 1 gcacgagggaacatttgcgc.....tcaaaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.ot.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.higo.hum.*
40: em.higo.mus.*
41: em.higo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1504.6	95.5	2156	8	AF276086	AF276086 Zea mays
2	1234	78.3	2247	8	AF276085	AF276085 Zea mays
3	987.4	62.7	2145	8	AK102647	AK102647 Oryza sat
4	987.4	62.7	2565	8	AK111501	AK111501 Oryza sat
5	979.4	62.1	1632	8	BT009111	BT009111 Triticum
6	577.8	36.7	1464	8	BT000625	BT000625 Arabidops
7	577.8	36.7	2090	8	AF250335	AF250335 Arabidops
8	577.8	36.7	2301	8	AF439826	AF439826 Arabidops
10	525.8	33.4	982	8	AY247268	AY247268 Medicago
11	241.6	15.3	1923	8	AK121122	AK121122 Triticum
12	135	8.6	92392	8	AC128647	AC128647 Oryza sat
13	135	8.6	145115	8	AC096690	AC096690 Oryza sat
14	87	5.5	107857	2	AC146552	AC146552 Medicago
15	77.2	4.9	110514	8	F7G19	AC000106 Sequence
16	73.6	4.7	139089	2	AC146571	AC146571 Medicago
17	55.8	3.5	90373	3	AC115579	AC115579 Dictyoste
18	52.2	3.3	945	11	CNS06ICE	AL400116 T7 end of
19	50.6	3.2	1905	6	AX482695	AX482695 Sequence
20	50.6	3.2	2082	3	AF094764	AF094764 Drosophil
21	50.6	3.2	2577	3	BT001510	BT001510 Drosophil
22	50.6	3.2	2684	3	AY166755	AY166755 Drosophil
23	50.6	3.2	2687	3	BT003249	BT003249 Drosophil
24	50.6	3.2	2871	3	AY166754	AY166754 Drosophil
25	50.6	3.2	3157	3	AY166758	AY166758 Drosophil
26	50.6	3.2	3199	3	AY166753	AY166753 Drosophil
27	50.6	3.2	3210	3	AY084138	AY084138 Drosophil
28	50.6	3.2	3318	3	AY166757	AY166757 Drosophil
29	50.6	3.2	3497	3	AF094763	AF094763 Drosophil
30	50.6	3.2	3538	3	BT010233	BT010233 Drosophil
31	50.6	3.2	3679	3	AY166756	AY166756 Drosophil
32	50.6	3.2	4387	3	BT003626	BT003626 Drosophil
33	50.6	3.2	4666	3	AY070541	AY070541 Drosophil
34	50.6	3.2	4783	3	AY166752	AY166752 Drosophil
35	50.6	3.2	80069	2	AC019671	AC019671 Drosophil
36	50.6	3.2	173634	3	AC009344	AC009344 Drosophil
37	50.6	3.2	195868	3	AC008308	AC008308 Drosophil
38	50.6	3.2	235707	3	AE003733	AE003733 Drosophil
39	49.6	3.1	969	6	AX595096	AX595096 Sequence
40	49.6	3.1	969	6	AX818518	AX818518 Sequence
41	49.6	3.1	969	6	AX829548	AX829548 Sequence
42	49.6	3.1	1434	8	YSCSNF4	M30470 Yeast (S.ce
43	49.6	3.1	1619	8	YSCCAT3	M21760 Yeast (S.ce
44	49.6	3.1	1818	8	SCYGL115W	Z72637 S.cerevisia
45	49.6	3.1	163443	2	AC006280	AC006280 Plasmodiu

ALIGNMENTS

RESULT 1
AF276086
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2156)
Lumbreras,V., Alba,M.M., Kleinow,T., Koncz,C. and Pages,M.
Domain fusion between SNF1-related kinase subunits during plant

AF276086
Zea mays protein kinase AKINbeta gamma-2 mRNA, complete cds.
2156 bp mRNA linear PLN 13-AUG-2002
AF276086
AF276086.1 GI:11139547
Zea mays
Zea mays
Zea mays

evolution
EMBO Rep. 2 (1), 55-60 (2001);
MEDLINE 21146516
PUBMED 11252725
REFERENCE 2 (bases 1 to 2156)
AUTHORS Lumbres, V. and Pages, M.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2000) Genetica Molecular, CID, CSIC, Jordi Girona,
18-26, Barcelona 08034, Spain
LOCATION/Qualifiers
source 1. 2156
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon-4577"
/dev_stage="seedling"
/note="under water stress"
456..1946
/note="Scf-4 like"
/codon_start=1
/product="protein kinase AKINbetagamma-2"
/protein_id="AA031752.1"
/db_xref="GI:11139548"
/translation="MESHGADSAHDAGAVSSGGATVPRFVMPYGGKRVFVSGSPT
RWSEHPMSPIEGCTVFOAICSLSPGHEIKFVDGEMHDERQPTISGEFGIVNTL
YLREYNQIMTLSPSPGPMNDVNFQRTVLSQDVTSEGLRVSEAAIQISR
CRYSEYLSLNTCYDLLPDSKVIALDNLVFKQSEHLLHSGQIPVAPLWDSFRQFVG
LLSPDLFILRLRETHGSLNLTBEQLETHLSAWKEAKROTGRNDRQWRPQOHLVHA
TPYSELDIAVLKQNGISIVPIYSSSDGSPFOLLHLASLSILKLCICRYFKNSTG
NLPILOPVCISPIGSPVVKPGLNLSPLMLPNSLSALANLVOAGVSSPIVDD
NDSLDYTSRSDITAKAKVYTHRLDENTHQAQLQGDANTPFQFNGRCQMCL
RSDPLKLMERLANPGRVRFIVEAGSKRVEHISLSIDIEFKLLSL"

ORIGIN

Query Match 95.5%; Score 1504.6; DB 8; Length 2156;
Best local Similarity 98.0%; Pred. No. 0;
Matches 1537; Conservative 2; Mismatches 24; Indels 8; Gaps 1;

QY	8	GGAACTTTCCCGATGCTCTCTATCGAAGGCTGCCACCTGTATTTTCAGGCTATTTTCAG	67
DB	596	GGAACTTTCCCGATGCTCTCTATCGAAGGCTGCCACCTGTATTTTCAGGCTATTTTCAG	655
QY	68	TCTGTCTCTCTGGGATTCACAGTACAAATTTCTTTGTGGACGGGAATGGCGCATGATGA	127
DB	656	TCTGTCTCTCTGGGATTCACAGTACAAATTTCTTTGTGGACGGGAATGGCGCATGATGA	715
QY	128	GGCTCAACCTACATCTCTGGGAGTTTGGCATAGTTTAAACACCTTTACTTTGACAGGGA	187
DB	716	GGCTCAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACCTTTACTTTGACAGGGA	775
QY	188	ATATAACCAATAAACCCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGAACATGGA	247
DB	776	ATATAACCAATAAACCCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGAACATGGA	835
QY	248	TGTGGATAATGAAATTTTCAACGTACGGTTAGCTGTACAGTGGCCACCGTTTCAGAAAG	307
DB	836	TGTGGATAATGAAATTTTCAACGTACGGTTAGCTGTACAGTGGCCACCGTTTCAGAAAG	895
QY	308	TACTCTGAGATTTTCAGAGCTGCATACAAATATCTAGGTTCGTCTTCTGATATCT	367
DB	896	TACTCTGAGATTTTCAGAGCTGCATACAAATATCTAGGTTCGTCTTCTGATATCT	955
QY	368	GAATTTGCATACATGCTATGATTTACTCCACAGATTCTGGCAAGGTTATTCGCCCTAGACAT	427
DB	956	GAGTTTGCATACATGCTATGATTTACTCCACAGATTCTGGCAAGGTTATTCGCCCTAGACAT	1015
QY	428	TAATTTACCTGTGAAGCAATTCATTCATATCTCCATGAAACAGGGGATTCCTGAGCTCC	487
DB	1016	TAATTTACCTGTGAAGCAATTCATTCATATCTCCATGAAACAGGGGATTCCTGAGCTCC	1075
QY	488	TCTCTGGGACTCATTCAGAGGTCATTTGTGGTCTCTCTAGCCCATTTGGATTTTCATCT	547
DB	1076	TCTCTGGGACTCATTCAGAGGTCATTTGTGGTCTCTCTAGCCCATTTGGATTTTCATCT	1135

QY	548	CATATTCGGGAGGCTAGAACTCATGGCTGAACTTTCAGACAGGACGCTTGAACACA	607
DB	1136	CATATTCGGGAGGCTAGAACTCATGGCTGAACTTTCAGACAGGACGCTTGAACACA	1195
QY	608	CACATATCTGCATGAAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATGTCAGTG	667
DB	1196	CACATATCTGCATGAAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATGTCAGTG	1255
QY	668	GGACCCGACACAGCATCTAGTGCATGCCACCCCTTATGAGTCTTGGAGGACATTCGAGT	727
DB	1256	GGACCCGACACAGCATCTAGTGCATGCCACCCCTTATGAGTCTTGGAGGACATTCGAGT	1315
QY	728	AAAGCTTTTGCAAAATGSCATTTCTACAGTGCACAGTATTTATTCATCATCATCAGATGG	787
DB	1316	AAAGCTTTTGCAAAATGSCATTTCTACAGTGCACAGTATTTATTCATCATCATCAGATGG	1375
QY	788	ATCATTTCCCGCAGTTATTGTCATCTTGCATCACCTTTCTGGAAATTTTGAATATTTTGTAG	847
DB	1376	ATCATTTCCCGCAGTTATTGTCATCTTGCATCACCTTTCTGGAAATTTTGAATATTTTGTAG	1435
QY	848	ATACTTCAAAAACCTCAACTGTAATTTGCTTCTGAAACCAACAGTGTGCTCCATTC	907
DB	1436	ATACTTCAAAAACCTCAACTGTAATTTGCTTCTGAAACCAACAGTGTGCTCCATTC	1495
QY	908	GCTGGGTTCTCTGGGTTCCGAAAATTTGGTGATCTCGAACAGTCTGCTTGGCTATGTTGCG	967
DB	1496	GCTGGGTTCTCTGGGTTCCGAAAATTTGGTGATCTCGAACAGTCTGCTTGGCTATGTTGCG	1555
QY	968	ACCTAATGCTCACTTAGCTCTGCCCTTAAACATGTTGGTTTCAAGCTGGAGTAAGCTCAAT	1027
DB	1556	ACCTAATGCTCACTTAGCTCTGCCCTTAAACATGTTGGTTTCAAGCTGGAGTAAGCTCAAT	1615
QY	1028	ACCAATTTGTCGATGACAAACGACTCCCTGCTTGACACTTCTCTAGAGTGCATCAGAC	1087
DB	1616	ACCAATTTGTCGATGACAAACGACTCCCTGCTTGACACTTCTCTAGAGTGCATCAGAC	1675
QY	1088	CCTAGCAAAAAGACAAGGCTTACACACATGTTGGCTGGATGAGATGACCATTCACAGGC	1147
DB	1676	CCTAGCAAAAAGACAAGGCTTACACACATGTTGGCTGGATGAGATGACCATTCACAGGC	1735
QY	1148	TTTTCAGCTTTGGAACAAGTCCCAATACACCTTTTGGATTCTTTAAACGCCACAGATGCCA	1207
DB	1736	TTTTCAGCTTTGGAACAAGTCCCAATACACCTTTTGGATTCTTTAAACGCCACAGATGCCA	1795
QY	1208	GATGTGCTCGGCTCTGATCTCTTGTGGAAGTGTGAGGCGACTGGCTAATCTCGGGT	1267
DB	1796	GATGTGCTCGGCTCTGATCTCTTGTGGAAGTGTGAGGCGACTGGCTAATCTCGGGT	1855
QY	1268	GGCGCGGCTGTTTCAATTTGGAAGCTGGGAGCAAAACGTTGTGGAGGCGCATCATCAAG	1327
DB	1856	GGCGCGGCTGTTTCAATTTGGAAGCTGGGAGCAAAACGTTGTGGAGGCGCATCATCAAG	1915
QY	1328	TGACATTTTCAAGTTCTTGTGAGCTTGTGAGGAGAGGAGCATGCGGAGTTTTTTTT	1387
DB	1916	TGATATTTTCAAGTTCTTGTGAGCTTGTGAGGAGAGGAGCATGCGGAGTTTTTTTT	1975
QY	1388	AACAGTAGCAACCTCTGGGCTTTGGGCTCCTTGGGATTAATTCGCTATCTCTCAGGACT	1447
DB	1976	AACAGTAGCAACCTCTGGGCTTTGGGCTCCTTGGGATTAATTCGCTATCTCTCAGGACT	2027
QY	1448	CAGGTTTCATGGTCTGTGAAAGTTCTCTTTTGTGTTTGGTTTTTACCATTTGGCCCTC	1507
DB	2028	CAGGTTTCATGGTCTGTGAAAGTTCTCTTTTGTGTTTGGTTTTTACCATTTGGCCCTC	2087
QY	1508	CCGTTTGTAAACAAACATACATCAGGTGATTTTATCTATTTGATTTTCAAAAAAAA	1567
DB	2088	CCGTTTGTAAACAAACATACATCAGGTGATTTTATTTATTTATTTGATTTTCTTTGAAAAA	2147
QY	1568	AAAAAATAA 1576	
DB	2148	AAAAAATAA 2156	

RESULT 2	AF276085	2247 bp	linear	PLN 13-AUG-2002
LOCUS	AF276085	2247 bp	linear	PLN 13-AUG-2002
DEFINITION	Zea mays protein kinase AKINbetagamma-1 mRNA, complete cds.			
ACCESSION	AF276085			
VERSION	AF276085.1	GI:11139545		
KEYWORDS				
SOURCE	Zea mays			
ORGANISM	Zea mays			
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	1 (bases 1 to 2247)			
TITLE	Lumbreras, V., Alba, X.M., Kleinow, T., Koncz, C. and Pages, M.			
JOURNAL	Domain fusion between SNF1-related kinase subunits during plant evolution			
MEDLINE	EMBO Rep. 2 (1), 55-60 (2001)			
PUBMED	21146516			
REFERENCE	1 (bases 1 to 2247)			
AUTHORS	Lumbreras, V. and Pages, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (07-JUN-2000) Genetica Molecular, CIB, CSIC, Jordi Girona, 18-26, Barcelona 08034, Spain			
FEATURES	Location/Qualifiers			
source	1..2247			
	/organism="Zea mays"			
	/mol_type="mRNA"			
	/db_xref="taxon:4577"			
	/dev_stage="seedling"			
	/note="under water stress"			
	496..1389			
	/note="Snf-4 like"			
	/ccdon_start=1			
	/product="protein kinase AKINbetagamma-1"			
	/protein_id="AA031751.1"			
	/db_xref="GI:11139546"			
	/translation="MFSSGADSDAGDVGWSSGVGATVPARFVMPYGGKRVFVSGSF TRWSHLFMSVPEGVKTFVQAIKSLPGHIEKFDVGEVHDERQPTISGEFQIVNT LYLTFEPIFNALLMPSPTGSRMMNDVNFCHTVLSQGTIPEGHVRVSEMAIQLS RCVSVEYLNLTCDLLPDSKVLALDINLPVKOSFHLLHEQGIPEVAPLMSRFGQV GLLSPLDFILILRELTGHSNLTDELTHTLSAKKEAKRQTCGRNDGQWRQHLVH ATPYSELSDIAVKLLNDISTVPIYSSSDGSPQLLHLASLSGLDKCPRIYFNKST GMLPLNGFVCSIPILGSPVWPKIPGPNFPLAMSLNLSALNKLVOAGVSSIPVD ENLSLDYRSRDIYTHALAKVYTHRLDEMAIHQALQGDQANTPFQFFNGORCOMC LRSDDLKVMERLANPVGVRVFEIVGAGSKRVEGIIISLSDIPKF-LSL"			
CDS				
	/product="protein kinase AKINbetagamma-1"			
	/protein_id="AA031751.1"			
	/db_xref="GI:11139546"			
	/translation="MFSSGADSDAGDVGWSSGVGATVPARFVMPYGGKRVFVSGSF TRWSHLFMSVPEGVKTFVQAIKSLPGHIEKFDVGEVHDERQPTISGEFQIVNT LYLTFEPIFNALLMPSPTGSRMMNDVNFCHTVLSQGTIPEGHVRVSEMAIQLS RCVSVEYLNLTCDLLPDSKVLALDINLPVKOSFHLLHEQGIPEVAPLMSRFGQV GLLSPLDFILILRELTGHSNLTDELTHTLSAKKEAKRQTCGRNDGQWRQHLVH ATPYSELSDIAVKLLNDISTVPIYSSSDGSPQLLHLASLSGLDKCPRIYFNKST GMLPLNGFVCSIPILGSPVWPKIPGPNFPLAMSLNLSALNKLVOAGVSSIPVD ENLSLDYRSRDIYTHALAKVYTHRLDEMAIHQALQGDQANTPFQFFNGORCOMC LRSDDLKVMERLANPVGVRVFEIVGAGSKRVEGIIISLSDIPKF-LSL"			
ORIGIN				
Query Match	78.3%	Score 1234	DB 8	Length 2247
Best Local Similarity	88.0%	Pred. No. 0	Mismatches 155	Indels 37
Matches 1414	Conservative 0	Gaps 5		
QY	8	GGAACTATTGCGGATGTCCTCTATCGAAGCGCTGCCCACTGATTTTCAGGCTATTTCAG 67		
Db	639	GGAACTATTGCGGATGTCCTCTGTCGAAGCGCTGCCCACTGATTTTCAGGCTATTTCAG 698		
QY	68	TCTGTCCTCGGATTCAGAGTCAAAATCTTTGTGACGGGATGAGGATGATGA 127		
Db	699	CTGTGTCCTCGGATTCAGAGTCAAAATCTTTGTGACGGGATGAGGATGATGA 758		
QY	128	GGGTCAACCTACCATATCTGGGGAGTTGGCATAGTTAAACACATTTACTTGAACAGGGA 187		
Db	759	GGCCAACTACTATATCTGGGGAGTTGGCATAGTTAAACACATTTACTTGAACAGGGA 818		
QY	188	ATATTAACAAATTAACACCTTATCAAGTCCCAAGCACACCTGGAACGAGATGAACATGGA 247		
Db	819	ATTTAATCAATAAATGTCCTTATTAATCCCAAGCACACCTGGAACGAGATGAACATGGA 878		
QY	248	TGTGGATATGAATAATTTTCAAGTACGCTTACGTTGTCAGATGCGACCGTTTCAGAGG 307		
Db	879	TGTGGATATGAATAATTTTCAACATACCGTTTACATTTGTCGGATGATGATTCAGAGG 938		
QY	308	TACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTTCTGAATATCT 367		

Db 2019 TTGTGATCAATAGTAAAGAAACAGCAGGCTATCTGAAGCAAGCTGCCATGATCTCCCA 2078
 Qy 1439 T-----CAGGACTCAGGCTCATGCTCTGTAAGTCTCTTT-- 1476
 Db 2079 TGGGCTTTGCTTTTGTGATGAATAGCTGTTGGAATTCATGCTCTCTAAAGTCTCTTTT 2138
 Qy 1477 ----TTTTGTTGCTTTTACCATTGTCCTCCCTGTTGTTTAAACAAAACATACATCNG 1532
 Db 2139 CCCTTGTGTTTTCGTTTAAACATTGTGCTCCCTGTTGTTTACCAAAAATATCTGCTG 2198
 Qy 1533 GTGATTTTAT--CTATTGATGTTTCAAAAAAATAAAAAA 1576
 Db 2199 ATTATTATTATTATTGTTGATGTTTCTCCAAAABAAAAA 2244

RESULT 3

AK102647
 LOCUS AK102647.1 GI:32987856
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033100L02, full insert sequence.
 ACCESSION AK102647
 VERSION AK102647.1
 KEYWORDS FUJ_CDNA; CAP trapper.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1
 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, X., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, K., Imamura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Iotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootani, K., Oka, Y., Saito, R., Sakai, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from Japonica rice
 Science 301 (5631), 376-379 (2003)
 MEDLINE 22752273
 PUBMED 12869764

TITLE

Japonica rice
 Science 301 (5631), 376-379 (2003)
 MEDLINE 22752273
 PUBMED 12869764
 2 (bases 1 to 2145)

REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirao, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, J., Iotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, M., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohtsuki, K., Oka, M., Ooka, H., Ootani, K., Ootani, R., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

TITLE

JOURNAL

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Teukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Iotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, N., Ootani, K., Oka, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.

FEATURES

Location/Qualifiers

1..2145
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J033100L02"

ORIGIN

Query Match 62.7%; Score 987.4; DB 8; Length 2145;
 Best Local Similarity 83.2%; Pred. No. 3.5e-262;
 Matches 1123; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

Qy 5 GAGGGAACATTTGCGGATGTCCTCATCGAAGGCTGCCCACTGTATTTCAGGCTATTG 64
 Db 200 GACAGAACATTTGCCAATGTCCTCTGTGAAGGTGTCCCACTGTTTTCAGGCCATATG 259
 Qy 65 CAGTCTCTCTCTGGGATTCACAGATCAAAATTCCTTTGTGGACGGGATGGCGCATGA 124
 Db 260 CAGTTGTGCTCCCTGGGATTCATCAGTACAGTTTGTGTGGATGGGAATGGCGCACGA 319
 Qy 125 TGACGGTCACCTACCATATCTGGGGAGTTTGGCATAGTTAAACACACTTTTACTTGACAAG 184
 Db 320 TGACGGCAACCTACCATAAACAGGAGACTATGGTGTGGTTAACTTTATGCTTGACTAG 379
 Qy 185 GGAATATAACCAATAAACACCTTTATCAAGTCCAAAGCACACCTCGAAGACGATGAACAT 244
 Db 380 GGACTTTGACCAGCAATAAACAATATTGAGCCCTAGTACACCTCGAAGTAGTAGTAACAT 439
 Qy 245 GGATGTGGATTAAGAAATTTCAACCTACGTTACGTTACGTTTCAGATGCCACCGTTTACA 304
 Db 440 GGATGTGGCAACGACAAATTTCAACGCTACTGTTTCTTTGTCGATGGCATTAATTCAGGA 499
 Qy 305 AGGTACTCTCAGAGCTTTTCAGAGGCTGCAATACAAATATCTAGGTGCTGTTCTGTAATA 364
 Db 500 AGTTCCTCAGGAATTTTCAGAGGAGCTATACAGATCTTAGGTGTCTGTAGGGGATTT 559
 Qy 365 TCTGAATTCGATACATGCTATGATTTTACTCCAGATTTCTGGCAAGGTTATTTGCCCTAGA 424

Db 560 TCTGAATGACAACTGGGTATGATTACTCCAGATTCTGGCAAGTCAATTCCTCTAGA 619
 QY 425 CATTAAATTAACCTGTGAAGCAATCAATTCATATTCCTCATGAACAGGGAATTCCTGTAGC 484
 Db 620 CATTAAATTCCTGTGAAGCAATTCCTCATATTCCTCATGAACAGGGAATTCCTGTAGC 679
 QY 485 TCCTCTCTGGGACTCATTCAGAGGTCAATTTGTGTCTCTCTAGCCCAATTCGATTTCAT 544
 Db 680 ACCCTATGGGATTCATTCAGGGCCAGTTTGTGGCCCTTTGAGCCCACTGGATTTCAT 739
 QY 545 ACTCATATTCGGGAGCTAGAACTCATGCTCGAACTTCGACAGAGACGCTTCAAAAC 604
 Db 740 ACTCATATTCGAGAGCTGGAACCTCATGCTCCAACTTCGACAGAGACGCTTCAAAAC 799
 QY 605 ACACACTATTCGATGTAAGAGAGCTTAAGGGCAACTTAATGGAAGAAATGATAGTCA 664
 Db 800 ACATACATATTCGCTGGTGAAGGGCCCAAGCGTCAAACTTATGCGCAGAAACGAGGGTTC 859
 QY 665 GTGGGACCGCAACAGCATCTAGTCATGCGACCCCTTATGAGTCTCTGAGGACATTCG 724
 Db 860 TTGGAGGGCAATCACCATTATGATTCATGCCACCCCTTATGATTCCTCGAGGAATTCG 919
 QY 725 AGTAAAGCTTTTGCMAAATGGCATTTCTACAGTGCAGTATTATTCATCATCATCAGA 784
 Db 920 TATGAAGTACTGCAAAACGGTCTTCTACCGTTCCAATATATGTTTCTCTCATCACCAGA 979
 QY 785 TGGATCATTCCTGGCAGTATTCGATCTGCATCATCTCTGGATCTTGGAAATGATTTG 844
 Db 980 TGGCTCATATCACAATTTGTCATCTGCTTCCCTTCTCGGAATTTTGAATGTATTTG 1039
 QY 845 TAGATACCTTCAAAACCTCAACTGGTAAATTTGCTTATCTGAAACCAACCACTGTGCTCCAT 904
 Db 1040 TAGATATTTTAAATTTCTCAAGTAAATTTACCTATTTTGAGCCCACTGTATGCACAT 1099
 QY 905 TCGGTGGGTTCTCGGTTCCGAAATTTGTGTATCTGAAACAGTCGTCATTCGCTAGCT 964
 Db 1100 TCCTCTGGGTAGCTGGGTTCCAAATTTGTGTATCTGATCTCTAATGCGGCTCCATTTGCTATGTT 1159
 QY 965 GGGACTATGCTCATCTAGCTGCTGCTTACATGTTGGTTCAGCTGGAGTAGCTC 1024
 Db 1160 GGGGCTCAACATCTCTTAGCGCTGCTGCTGAAATTTGCTGGTTCAAGCTGGTGTAGCTC 1219
 QY 1025 AATACCAATTTGGATGACAAACGACTCCCTGCTGTGACACTTACTCTAGAGTGCATCATC 1084
 Db 1220 AATACCAATTTGGATGACAAATGACTCATCTGCTGTGACACATATCCAGAGTGCATCATC 1279
 QY 1085 AGCCCTAGCAAAAGCAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCA 1144
 Db 1280 AGCTCTTGTCAAAAGCAAGGTTGTACACACATTCGCTAGATGAGATGACCATTCATCA 1339
 QY 1145 GCGTTTGCAGCTTGGACAGATGCCAATACACCTTTTGGATCTTTAAAGCCGACAGATG 1204
 Db 1340 GGCCTGCGAGCTTGGACAGATGCAATTCGCCCTTTGGATCTTTAAAGCCGACAGATG 1399
 QY 1205 CCAGATGTCCTCCGCTCTGATCTCTTGTGTAGGTGATGGAGCTGCTGCTAATCCTGG 1264
 Db 1400 TCAGATGTCCTCCGCTCTGATCTCTTGTGTAGGTGATGGAGCTGCTAATCCTGG 1459
 QY 1265 GTGGGCGGGGTTCATTTGTGGAAGCTGGAGCAAAACGCTGTGGAGGGCATCTATCACT 1324
 Db 1460 GTGGGCGGGGTTCATTTGTGGAAGCTGGAGCAAAACGCTGTGGAGGGCATCTATCACT 1519
 QY 1325 AAGTGACATTTTCAAGTCTTCTGCTGAGCT 1353
 Db 1520 GAGTGTATTTTCAAGTCTTCTGCTGAGCT 1548

RESULT 4

AK111501

LOCUS

DEFINITION

Accession

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

AK111501.1 GI:37988164
 FLI_CDNA; CAP trapper.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Sukayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1
 The Rice Full-length cDNA Consortium, National Institute of
 Agricultural Sciences Rice Full-length cDNA Project Team,
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group: Ohtsuki, K., Murakami, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Narioka, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, J.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M. and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice
 Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Satoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y.,
 Nakamura, M., Nami, T., Narioka, R., Niikura, J., Niikura, Y.,
 Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M.,
 Ooka, H., Osato, N., Ota, Y., Ohtsuki, K., Ohtsuki, K., Sakai, C.,
 Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
 Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y.,
 Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,
 Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K.,
 Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,
 Yokomizo, S. and Yoshimura, A.
 Rice full-length cDNA
 Unpublished
 3 (bases 1 to 2565)
 Kikuchi, S.
 Direct Submission
 Submitted (12-SEP-2002) Shoshi Kikuchi, National Institute of
 Agricultural Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 32K full-length cDNA clones from japonica
 rice.
 URL : <http://cdna01.dna.affrc.go.jp/cdna/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,
 Yamamoto, M. and Nakahama, Y.
 FAIS Genome Sequencing & Analysis Group: Ohtsuki, J., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M.,
 Kobayashi, M., Kodama, T., Kurcsaki, T., Kusumegi, T., Lu, M.,

Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niiikura, J., Oka, M.,
 Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,
 Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Haragaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
 Ota, Y., Saichoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M.,
 Tagami-Takeda, Y., Tagawa, A., Takahashi, E.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
 Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers
 1. 2565
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J01300N04"

ORIGIN

Query Match 62.7%; Score 987.4; DB 8; Length 2565;
 Best Local Similarity 83.2%; Pred. No. 3.4e-262;
 Matches 1123; Conservative 0; Mismatches 226; Indels 0; Gaps 0;
 5 GAGGAAACATTTGCCGATGCTCTCTGTTGAAGTGTGTCACCGCTTTTTCAGGCAATG 64
 610 GACAGAAACATTTGCCGATGCTCTCTGTTGAAGTGTGTCACCGCTTTTTCAGGCAATG 669
 65 CAGTCTCTCTCTGGGATTCAGAGTACAAATCTTCTGTCAGCGGGAATGCGGCTGA 124
 670 CAGTTTGTCCCTGGGATTCATCAGTACAGTTTGTGTTGAGTGGGAAATGCGGCGCA 729
 125 TGAGCGTCAACTACCATATCTGGGGAGTTTGGCATAGTTTAAACACATTTTACTTGACAAG 184
 730 TGAGCGCAACTACCATATACAGGAGCTATGTTGTTTAAACATTTTATGCTTGACTAG 789
 185 GGAATATAACCAATAACACTTATTAAGTCCAGGACACCTGGAAGCAGATGAACAT 244
 790 GGAATTTGACAGATAAATAAATATTGAGCGCTAGTACACCTGGAAGTGAAGTGAACAT 849
 245 GGATGTGATATGAAATTTTCAACGTACGTTTACGTTTGTGATGTCAGATGCAACGTTTTCAGA 304
 850 GGATGTGCAACAGCAATTTTCAACGATCTGTTTCTTGTCTGATGTCATATTCAGGA 909
 305 AGGTACTCTGAGAGTTTCAGAGGCTGCAATATCAATATCTAGGTGTCTGTTTCTGAATA 364
 910 AGGTCTCTCAGAGATTTTCAGAGGCGAGCTATACAGATCTCTAGGTGTCTGTTAGCGGATTT 969
 365 TCTGAATTCATACATCTGATTTACTCCGATCTCGGAGTTTATGCGCTAGA 424
 970 TCTGAATTCGCAAACTGGGTATGATTTACTCCGATCTCGGAGTTTATGCTCTAGA 1029
 425 CATTAATTTACTCTGGAAGCAATCATTCATATCTCCATGAACAGGGGATCTCTGTAGC 484
 1030 CGTTAAATTCCTGTGAAGCAATCTTTTCAATATCTTCAATGAAGGGAATTCCTGTGCG 1089
 485 TCCTCTCTGGAGCTCATTCAGAGTCAATTTGTTGTTCTCTCCATGCCCATTTGATTTTCAAT 544
 1090 ACTCTATGGGATTCATTCAGGGGCGAGTTTGTGCGCTTTTGAAGCCACTGGATTTTAT 1149
 545 ACTCATATTCGGGAGCTAGAACTCATGCTCGAATTCAGAGGAGGAGCTTGAAAC 604
 1150 ACTCATATTCGAGAGCTGGAACATCTGCTCCCAATCTGACAGAGGAGGAGCTTGAAAC 1209
 605 ACACACTATCTCGATGGAAGAGGCTTAAGCGCAAACTAATGAAGAAATGATGATCA 664

1210 ACATCTATATCTGCTGGAGGAGGCGCAAGGTCACAACTTATGCCAGAAACGAGGGTTC 1269
 665 GTGGGAGCCGCAACAGCATCTAGTGCATATGCCACCCCTTATGAGTCCCTTGAGGACATTCG 724
 1270 TTGGAGGCAAAATCACCAATTTAGTTTCAATGCGACCCCTTATGAATCCCTTGAGGAAATTCG 1329
 725 AGTAAAGCTTTTGCAGAAATGSCATTTCTACAGTGCAGTATTTATTTATCATCATCATCAGA 784
 1330 TATGAAGATATCTGCAAAACGCTGTTCTTACCGTTCATATATGTTTCTCATCATCAGAGA 1389
 785 TGGATCATTTCCGCGAGTTATTTGATCTTTGATCATCTTTCTGGAATTTTGAATATGATTTG 844
 1390 TGGCTCATATCCCAATTTGTCATCTTCTTCCCTTCTGGAATTTTGAATATGATTTG 1449
 845 TAGATCATCTCAAAATCTCAACTGCTGTAATTTGCTTATTTCTGAAACCAACAGTGTCTCCAT 904
 1450 TAGATATTTTAAATAATTTCTCAAGGTAATTTACCTATTTTGGAGCAACCTGTATGCCAAT 1509
 905 TCCGCTGGGTTCTCGGGTTCCGAAATTTGGTGAATCTGAACAGTCTGTCATTTGGCTATGTT 964
 1510 TCCTCTGGGTACCTGGGTTCCAAATTTGGTGAATCTTAAATGCTGCTTCAAGTGGTGTGAGCTC 1569
 965 GCGACCTAATGCTCCTACCTAGTCTCTGCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
 1570 GCGGCTTAACACATCTCTTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629
 1025 AATACCAATTTGGTGAATGACAGTCCCTGCTTGAACATTTACTTACTAGAGTGACATCAC 1084
 1630 AATACCAATTTGGTGAATGACAGTCCCTGCTTGAACATTTACTTACTAGAGTGACATCAC 1689
 1085 AGCCTTAGCAAAAGCAAGGCTTACACACATTTTGGCTGGATGAGATGACCATTCACCA 1144
 1690 AGCTTTGCAAAAGCAAGGCTTACACACATTTTGGCTGGATGAGATGACCATTCACCA 1749
 1145 GCGTTGCGCTTGGCAAGATGCAATATACCTTTTGGATTTCTTTAAGCGGCAAGATG 1204
 1750 GCGTCTGCGCTTGGCAAGATGCAATATACCTTTTGGATTTCTTTAAGCGGCAAGATG 1809
 1205 CCAGATGCTCCGCTGCTGATCTTCTGCTGAAGTGTGAGCGGCTGCTGCTGCTGCTGCTGCTG 1264
 1810 TCAGATGCTCTCGGCTGCTGATCTTCTGCTGAAGTGTGAGCGGCTGCTGCTGCTGCTGCTG 1869
 1265 GGTGCGGCGGTGTTTCAATTTGGAAGCTGGGAGCAACAGTGTGAGGCGCATCATATCACT 1324
 1870 GGTGCGGCGGTGTTTCAATTTGGAAGCTGGGAGCAACAGTGTGAGGCGCATCATATCACT 1929
 1325 AAGTGAATTTTCAAGTCTTCTGCTGAGCT 1353
 1930 GAGTGAATTTTCAAGTCTTCTGCTGAGCT 1958

RESULT 5

BT009111 1632 bp mRNA linear PLN 20-JUN-2003
 LOCUS Triticum aestivum clone wll.pk0008.b3:fls, full insert mRNA
 DEFINITION
 ACCESSION BT009111
 VERSION BT009111.1 GI:32128662
 KEYWORDS FLI_CDNA.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 1632)
 AUTHORS Tingey, S.W., Walters, P., Powell, W., Dolan, M., Miao, G.-H.,
 Caraher, N.R., Hanafey, M.K. and Hainey, C.F.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. Dupont de Nemours and
 Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA
 FEATURES
 Location/Qualifiers
 1. 1632
 source

/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wll.pk0008.b3.fis"

ORIGIN

Query Match 62.1%; Score 979.4; DB 8; Length 1632;
Best Local Similarity 83.5%; Pred. No. 5.7e-260;
Matches 1125; Conservative 0; Mismatches 216; Indels 6; Gaps 1;
QY 12 CATTTGCCGATGTCCTTATCGAAGGCTGCCCCCACTGTATTTTCAGGCTATTTGCGAGTCTG 71
DB 2 CACGAGCGGATGTCCTCCAGTCGAAGGTTGCCCACTGTATTTTCAGGCTATTTGCGAATTA 61
QY 72 TCTCTCGGATTCACGAGTACAAATCTTTTGGACGGGGAATGCGGCATGATGAGCGT 131
DB 62 CTTCCAGGGAATTTATCAGTACAAGTTCAATGTGGACGGGCACTGGAGGCATGACAGGGG 121
QY 132 CAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACACTTTTACTTGAAGGGAATAT 191
DB 122 CAACTACTATAACTGGAGATATGGGTGGTAAACACTTTTATCTTGACAAGGGAATTT 181
QY 192 AACCAATAAACCCTTATCAAGTCAAGCAACCTGGAAGCAGGATGAACATGATGATG 251
DB 182 GACCACATAAATATCTACTGAGCCCCACTACACCTGGGAGCAGG-----ATGATGTG 235
QY 252 GATATGAAAAATTTCAAGTACGGTTTACGTTGTGTCAGATGGCACCGTTTTCAGAAGGTACT 311
DB 236 GACGTGACAGTTTCAAGCAATGGGTTGCTGTGCGAGTGGGCCCTTCAGGAAGTCTCT 295
QY 312 CTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTTCTGAAATATCTGAAT 371
DB 296 CCAAGAACTTCAGAGGCTGCTATACAGATCTCTAGGTGTCGTGTTCTGAGTATCTGAAT 355
QY 372 TGGATATGATGATGATTTACTCCAGATTTCTGCAAGTTTATGCGCCCTAGACATTAT 431
DB 356 GCGCATACAGGCTATGACCTACTACAGATTTCTGGAAGGTCAATGCTCTGACACATTAT 415
QY 432 TACCTGTGAAGCAATCATCTCCATATTCCTATGAACAGGGATTCCTGTAGTCTCTCTC 491
DB 416 TTACTGTGAAGCAATCTTCCATATTTCCATGTAACAGGGATTCCTGTGCTCTCTCTG 475
QY 492 TGGGACTCATTCAGAGGTCATTTGTTGTCCTCTTACGCCATGCGATTTCTATCTATCA 551
DB 476 TGGGATTCATTCAGGGTCAAGTTGTTGGCTCTCTGAGCCCACTGATTTTATCTATA 535
QY 552 TTGCGGGACTAGAACTCATGCTCGAATCTTGACAGAGAGCAGCTTGAACACACACT 611
DB 536 TTGAGAGAGCTGGAACCTCATGGCTCAAACTTGACAGAGGAAACAGCTTGAACACACACT 595
QY 612 ATATCTGATGAAAGAGCTAAGCGGCAAACTAATGGAAGAAATCATAGTCAAGTGGCGA 671
DB 596 ATATCTGCTGGAAGAGCTAAGCGGCAAACTAATGGAAGAAATCATAGTCAAGTGGCGA 655
QY 672 CCGGAACAGCTCTAGTCAATGCAACCCCTTATGATCTCTTGAAGCACTTGAAGTAAAG 731
DB 656 TCAAAATCAGCATCTAGTCAATGCAACCCCTTATGATCTCTTGAAGCACTTGAAGTAAAG 715
QY 732 CTTTTCGAAATGGCAATTTCTACAGTGGCCAGTATTTTATTCATCATCATGATGATCA 791
DB 716 ATACTGAAATGGCAATTTCTACAGTGGCCAGTATTTTATTCATCATCATGATGATCA 775
QY 792 TTCCCGCAGTATTTGATCTTGCATCACTTTCTGGAATTTTGAATATGATTTGTAGATAC 851
DB 776 TTTCGCACTGTTGATCTTGCATCTCTTTCAGGAAATTTTGAATGATCTGATGATAC 835
QY 852 TTCAAAACTCACTGGTAAATTTGCTATTTGCAACCAACAGTGTGCTCCATTCGGCTG 911
DB 836 TTCAGAACTCACTGGTAAATTTGCTATTTGCAACCAACAGTGTGCTCCATTCGGCTG 895
QY 912 GCTTCTGGTTCGGAATTTGGTATCTGACAGTGGCTCCATTCGCTATGTCGACCT 971
DB 896 GGTACTCTGGTTCGGAATTTGGTAAACCAATTTGGTCAATTCGCTATGCTGATGTCGGGCT 955

QY 972 AATGCTCACTTAGCTCTGCCCTTAACATGTTGGTTCAAGCTGGAGTAAAGCTCAATACCA 1031
DB 956 AATACATCTTTAGCTCTGCCCTTAACATGTTGGTTCAAGCTGGAGTAAAGCTCAATACCC 1015
QY 1032 ATTGTGATGACAAAGCACTCCCTGCTTGACACTTACTCTAGAAAGTGAACATCAAGCCCTA 1091
DB 1016 ATTGTGATGATAAAGCACTCGCTGATCGACACATATCTCCAGAAAGTGAACATCAAGCTCTA 1075
QY 1092 GCAAAAGCAAGGCTCTACACATGTTGGCTGGATGAGATGACCATTCACAGGCTTTG 1151
DB 1076 GCAAAAGCAAGGCTCTACACATGTTGGCTGGATGAGATGACCATTCACAGGCTTTG 1135
QY 1152 CAGCTTGGCAAGATGCCAATACACCTTTTGAATTTCTTAAACGCCAGAGATGCCAGATG 1211
DB 1136 CAGCTCGGCAAGACCGGAATTCACCTTTTGAATTTCTTAAACGCCAGAGATGCCAGATG 1195
QY 1212 TGCCTCCGCTGATCTTGTGTAAGGTGATGAGGAGCTGCTGCTAATCTCTGCGGTGCGG 1271
DB 1196 TGTCTCCAGCTCTGACCTTTGCTGAAAGTTATGAGAGATTTGCTAATCTCTGCGGTGCGT 1255
QY 1272 CGGGTCTTCATTTGGAAGCTGGGACAAACGTTGAGGGCATCATCTACTAAGTGAC 1331
DB 1256 CCGGTGTTTCTGTTGAGGCTGGCAAGCGAGTGAAGGCGTAATATCTGCTGAGCGAC 1315
QY 1332 ATTTTCAAGTTCTTCTGCTGAGCTTTGTA 1358
DB 1316 ATATTCAGTTGCTGCTGAGCTAGCGA 1342

RESULT 6

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

REFERENCE

TITLE

JOURNAL

AUTHORS

REFERENCE

TITLE

JOURNAL

AUTHORS

REFERENCE

TITLE

JOURNAL

AUTHORS

REFERENCE

TITLE

JOURNAL

AUTHORS

REFERENCE

TITLE

JOURNAL

AUTHORS

REFERENCE

TITLE

JOURNAL

AUTHORS

REFERENCE

TITLE

JOURNAL

AUTHORS

REFERENCE

BT000625 1464 bp mRNA linear PLN 25-SEP-2002
Arabidopsis thaliana putative activator subunit of SNF1-related
protein kinase SNF4 (At1g09020/F7619_11) mRNA, complete cds.

BT000625

BT000625.1 GI:23308442

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 1464)

Chen, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,

Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y.,

Huan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,

Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,

Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,

Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,

Theologis, A. and Ecker, J.R.

Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1464)

Chen, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,

Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y.,

Huan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,

Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,

Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,

Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,

Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (25-SEP-2002) Salk Institute Genomic Analysis Laboratory

(SIGHALL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

Riken Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN

Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, POEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Shin, P., Ban, J., Bowser, L., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers
 1. 1464
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="U14970"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51"
 1. 1464
 /note="putative activator subunit of SNF1-related protein kinase SNF4"
 /codon_start=1
 /product="At1g09020/p7g19_11"
 /protein_id="AA18191.1"
 /db_xref="GI:23308443"
 /translations="MFGSTLDSRGSAAAGSGLLTPFRVPMYGGRRVFLSGSTFRMT
 EHPVMPGPEPTVFOVNCITGHOYKFDVGEWNRHDEHPQFVSGGVGVNTFIT
 GPDMVPAEGSPVNDVDFVLTADPOBAVRMSGVLELSRHRISVLLSTR
 TAYELLPSKSVTALDNLVYKQAFHLYEGLPLAPLWDPGQFVGLPDLDFILL
 LRSLGTHGSLTBEETHLTAWKGKHHISQYDGRPRPLVQVGFVPLNKKDV
 ALXILQNKVAAPVIVSSLDGSPQLHLASGLKLCICRPRSSSLPILQOPI
 CSPLGTWPKRIGESSKPLATLPRFASGLSALALVQAEVSIIPVDDNDLSDIYS
 RSDITLAKQAKAQLHLDMTVHQALQGDASPPYGIENQGRCHMLRSLSLVKVM
 ERLANPQVRRLVIVEAGSKRVEGLISLSLDFVQFLGL"

CDS

1. 1464
 /note="putative activator subunit of SNF1-related protein kinase SNF4"
 /product="At1g09020/p7g19_11"
 /protein_id="AA18191.1"
 /db_xref="GI:23308443"
 /translations="MFGSTLDSRGSAAAGSGLLTPFRVPMYGGRRVFLSGSTFRMT
 EHPVMPGPEPTVFOVNCITGHOYKFDVGEWNRHDEHPQFVSGGVGVNTFIT
 GPDMVPAEGSPVNDVDFVLTADPOBAVRMSGVLELSRHRISVLLSTR
 TAYELLPSKSVTALDNLVYKQAFHLYEGLPLAPLWDPGQFVGLPDLDFILL
 LRSLGTHGSLTBEETHLTAWKGKHHISQYDGRPRPLVQVGFVPLNKKDV
 ALXILQNKVAAPVIVSSLDGSPQLHLASGLKLCICRPRSSSLPILQOPI
 CSPLGTWPKRIGESSKPLATLPRFASGLSALALVQAEVSIIPVDDNDLSDIYS
 RSDITLAKQAKAQLHLDMTVHQALQGDASPPYGIENQGRCHMLRSLSLVKVM
 ERLANPQVRRLVIVEAGSKRVEGLISLSLDFVQFLGL"

ORIGIN

Query Match 36.7%; Score 577.8; DB 8; Length 1464;
 Best Local Similarity 66.4%; Pred. No. 1.1e-148;
 Matches 901; Conservative 0; Mismatches 432; Indels 24; Gaps 4;
 5 GAGGACATTTCCGATGCTCTCTATCGAAGGCTGCCACCTGCTATTTCAGGCTATTG 64
 129 GACAGAGCATGTGCCAATGTCACCGCTCGAGGGCTGCCCTACTCTTTTCAAGTCATTG 188
 65 CAGTCTGCTCTCTGCGGATTCACAGTACAAATCTTTGTGGAACGGGAATGGCGGACGA 124
 189 CAACTTGACCGCAGGATATCATCATATAGTTTGTGATGGGAATGGCGGACGA 248
 125 TGAGGCTCAACCTACCATATCTCGGGAGTTTGGCATAGTTTAAACACCTTTACTTGACAAG 184
 249 TGAGCACCAACCATTTGTAAAGCGGAATGGTGGAGTAGTGAATACAAATATTATTAACGG 308
 185 GGAATATACCAATAAACACCTTATCAAGTCCAGCACACCTCGGAAGCAGATGAACAT 244
 309 GCCAGATA---TGGTCTCTGCTGTTTACGCCAGACACTGG---GCCGCTCGAATAT 362
 245 GGATGTGGAATAAGAAATTTTCAACGCTACGGTTACGTTGTGATGATGGCAGCCGTTTCAGA 304
 363 GGATGTGATGATGCTCTCTCGCA-----ACGGCTGACCCGCTCCACGA 407
 305 AGGTACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTGCTGTTCTGAATA 364
 408 AGCTGTCTTAGGATGTGGGGGTGTGATTTGGAGCTTCTCGTCAACCGCATATCGGTTT 467
 365 TCTGAATTTGCATACATGCTATGATTTTACTCCAGATTCTGGCAAGGTATTGCGCCTAGA 424
 468 ATTGTCAACCCGCACTGCATATGAGTGTGCTCCCAAGATCGGCAAGGTATTGATGATGGA 527
 425 CATTAATTTACGTGTGAAGCAATCAATTCATATTCCTCATGAACAGGGGATTCGTGTAGC 484

Db 528 TGTATTTCCAGTGAAGCAAGATTCATATATCTCTATGAGCAGGAATCCCTTGGC 587
 Qy 485 TCCTCTCGGATCATATTCAGAGGTCAATTTGTGTCTCTCTAGCCCATTTGGATTTCAT 544
 Db 588 TCCTCTTTGGGACTTTTGGAAAAGGCCAATTTGTGTGAGTTCTTGGTCCACTAGACTTCAT 647
 Qy 545 ACTCATATTCGGGGAGCTAGAAACTCATGCTCGAAGTTCAGCAAGAGAGAGCTTGAAC 604
 Db 648 TCTAATACTCAGAGAGCTTGGAACTCATGATCAACTTACAGAGGAAGAGCTTGAAC 707
 Qy 605 ACACACTATATCTGCATGGAAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATAGTCA 664
 Db 708 GCACAAATAGCAGCTGGAAAGAGGAGGCTCATATTAGCCGCAATATGATGAAG 767
 Qy 665 GTGGCAGACCGAAGC---AGCATCTAGTGCATGCGCAGCTTATGAGTCTTGGAGGACAT 721
 Db 768 TGGGAGACCATATCTTAGGCGCACTTGTTCAGGTTGGACCTTATGATTAATCTGAAGACGT 827
 Qy 722 TGCAGTAAAGCTTTGCAAAATGCGCATTTCTACAGTGCAGTATTTATTTATTCATCATCATC 781
 Db 828 TGCCTGAAAATTTTGCAAAACAGGTGGCAGCTGTTCCGGTATATATTTCTTTTGCA 887
 Qy 782 AGATGATCATTTCCCGCAGTTATTGCACTTTTGCATCATCTTTCTGGAATTTTGAATGAT 841
 Db 888 GGATGGTTTCAATATCCGAGTTACTGCACTTCTGCTTCATCAGGCAATTTAAATGAT 947
 Qy 842 TTGTAGATATCTTCAAAATCAACTGTGTAATTTGCCCTATTCTGAAACCAACAGCTGTGCTC 901
 Db 948 ATGCAGATATCTTGAACATTCGCTAGCTCTTTACCTATCTCTTCAGAACCCATTTGTTTC 1007
 Qy 902 CATTCCGCTCGGTTCTCGGGTTCGAAATTTGGTGTCTGAAACAGTGTCTCCATTTGGCTAT 961
 Db 1008 AATTCCCTCGGTACGTGGTCTCTAGAACTCGGAAATCAAGTACAAACCTCTCGCTAC 1067
 Qy 962 GTTCCGACCTTAATSCCTCATCTAGCTCTGCGCTTAAACATGTTGGTTCAAGCTGGAGTAAG 1021
 Db 1068 ATTGAGACCTCACGCTCTTTGGGTTCTCGGCTCGCATTTATTTCAAGCTGAAGTCAG 1127
 Qy 1022 CTCAATACCAATTTGTGATGACAGCTCTCTGCTTTGACACTTACTCTGAGAGTGCAT 1081
 Db 1128 TTCAATTTCCGCTAGTGGATGACAAACGACTCGCTTTTGGATATATCTCTCGAAGTATAT 1187
 Qy 1082 CACAGCCCTAGCAAAAGACAGGCTTACACACATGTTTCGGCTGGATGAGATGACCATCA 1141
 Db 1188 AACTGCTCTGCTTAAAGCAAGGATACGACAGATTCATCTTGATGACATGACAGTTCA 1247
 Qy 1142 CCAGGCTTTGCGAGCTTGGACAAGATGCCAATACACCTTTTGGATTCTTTAAGGCCAGAG 1201
 Db 1248 CCAGGCTTACAGTTGGGGCAAGATGCGAGCCGCTTATGGAATCTTCAACGGGCGAG 1307
 Qy 1202 ATGCCAGATGCTCGCTCGGTCTGATCTTTGCTGAAGTATGAGGCGACTGGGCTAATCC 1261
 Db 1308 ATGTCACTAGTCTGCTCGCTCAGACTCTCTGTGAAAGTATGAGCGGTTGGCGAATCC 1367
 Qy 1262 TGGGCTGCGCGGGTGTTCATTGTGAAAGCTGGGAGCAAAAGCTGTGAGGAGCATCATATC 1321
 Db 1368 AGGTGTAAGGAGGCTGGTGTATGTGGAAGCAGGAGCAAAAGCTGTGAAAGTATCATTTTC 1427
 Qy 1322 ACTAAGTGACATTTCAAGTTCTTGTGAGCTTGTGA 1358
 Db 1428 ACTGAGCGATTTTCCAAATTTCTGCTCGGCTTTGA 1464

RESULT 7

AF250335 2090 bp mRNA linear PLN 06-SEP-2000
 LOCUS Arabidopsis thaliana putative activator subunit of SNF1-related
 DEFINITION protein kinase SNF4 mRNA, complete cds.
 ACCESSION AF250335
 VERSION AF250335.1 GI:9965728
 KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM	Arabidopsis thaliana	Db	1084	ATTGTCAACCCGCACTGTCATATGAGCTGCTCCAGAAATCGGCAAGGTATTATTCATTTGA 1143
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 2090)	Qy	425	CATTAACTTACCTGGAAGCAATCATTCATATTTCTCCAGAAAGAGGATTCCTGAGC 484
AUTHORS	Kleinow,T., Bhalarao,R., Breuer,F., Umeda,M., Salchert,K. and Koncz,C.	Db	1144	TGTTAATTTGGCAGTGAAGCAATTCATATATCTATGACGAGGAATCCCTTTGGC 1203
TITLE	Functional identification of an Arabidopsis snf4 ortholog by screening for heterologous multicopy suppressors of snf4 deficiency in yeast	Qy	485	TCCTCTCTGGGACTCATTTGAGAGTCAATTTGTTGGTCTCTCTTAGCCCATTTGATTTCA 544
JOURNAL	Plant J. 23 (1), 115-122 (2000)	Db	1204	TCCTCTTTGGGACTTTGGAAGAGCCAAATTTGTTGGAGTTCTTTGGTCCACTAGACTTCAT 1263
MEDLINE	20387008	Qy	545	ACTCATATTTGGGAGCTAGAACTCATGCTGAACTTGAACCTTGCACAGAGAGAGTGAAC 604
PUBMED	10929106	Db	1264	TCTAATATCTGAGAGAGCTTGGAACTCATGATCAAACTTGCACAGAGAGAGTGGAGC 1323
REFERENCE	2 (bases 1 to 2090)	Qy	605	ACACATATATCTGCATGGAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATCTCA 664
AUTHORS	Kleinow,T., Bhalarao,R. and Koncz,C.	Db	1324	GCACAATAGCAGCTTGAAGAGGAGGAGGCTATATTAGCCGCAATATGATGGAAG 1383
TITLE	Direct Submission	Qy	665	GTGGGACCGCAAC---AGCATCTAGTGCATGCCACCCCTTATGAGTCCTTTGAGGGACAT 721
JOURNAL	Submitted (29-MAR-2000) Genetic Principles of Plant Breeding, Max-Planck-Institute for Plant Breeding, Carl-von-Linne-Weg 10, Koeln 50829, Germany	Db	1384	TGGGAGACCATATCTTAGGCCACTTGTTCAGGTTGGACCCCTATGATAATCTGAAAGAGCT 1443
FEATURES	Location/Qualifiers	Qy	722	TGCAGTAAGCTTTTGCAGAAATGGCAATTTCTACAGTGCAGTATTTATTCATCATCATC 781
source	1..2090	Db	1444	TGCCCTGAAAAATTTTGCAGAAACAAAGGTGGCAGCTGTTCCGTTTATATATATTTCTTTGCA 1503
CDS	/note="contains 4 cyst-athione-beta-synthase (CBS) domains; heterologous multicopy suppressor of snf4 deficiency in yeast"	Qy	782	AGATGATCATTTCCCGCAGTTATTTGCATCTTGCATCACTTTCTGGAAATTTGAAATGAT 841
	/codon_start=1	Db	1504	GGATGTTTANATCCGAGTTACTTGCATCTTGTCTTCTGCTATCAGGCATATATAAATGAT 1563
	/product="putative activator subunit of SNF1-related protein kinase SNF4"	Qy	842	TTGTAGATATCTTCAAAACTCAACTGCTAATTTTGCCTTATTTCTGAAACCAACAGTGTCTC 901
	/protein_id="AAG10141.1"	Db	1564	ATGCAGATATCTTAGACATTTGCTAGCTCTTTTACCTATCTCTTTCAGCAACCCATTTCTTC 1623
	/db_xref="GI:9365729"	Qy	902	CATTCCGCTGGGTTCTCGGTCGGAATAATGTTGATCTGAAACAGTCTGTCATTTGGTAT 961
	/translation="MVPAGFSPETIGRSNMVDVFLRTADPSQERAVPMRGVDLELS RHRSVLTSTRAYELLPSGKVIADVNLVVKQAFHLYEQIPLAWDFGKGQFV GVLPDLFTLIURELTHGSNLTETEELETHHTAAWKEGKAHISROYDGSGRYPRLV OXGPDNLKDLVAKLQNKVAAVPVTVSSLODGSYFOLLALHASLSGLKICRYPHRS VSSPLIQPICSIPLGTWVPRIGSSSKPLATLRPHASLSGALALLVQAEVSSIPVY DNDLSLDIYSIDITALAKAYAIQHDLMFVHQALGQDASPPYGFNGQRCMH CLRSLSLVKMERLANPVRRLVIVKAGSRVEGHIISLSDVFQFLGL"	Db	1624	AATTTCCCTGGGTACGTGGGTCCTTAGATCGGAGAAATCAAGTAGCAAACTCTTTGCTAC 1693
		Qy	962	GTTCGACCTAATGCTCTACCTAGCTCTGCCCTTAACTGTTTCAAGCTGGAGTAAG 1021
ORIGIN		Db	1684	ATTGAGACCTCAGCCCTCTTTGGGCTCTGGCTCGCATTAATTAATTTCAAGCTGAGTCAG 1743
Query Match	36.7%; Score 577.8; DB 8; Length 2090;	Qy	1022	CTCAATACCAATTTGSGATGACAAACGACTCCCTCGCTTTGACACTTACTCTAAGAGTGACAT 1081
Best Local Similarity	66.4%; Pred. No. 1.1e-148;	Db	1744	TTCAATTTCCGTTAGTGGATGACAAACGACTCGCTTATTGACATATACTCTCGAAGTGATAT 1803
Matches	901; Conservative 0; Mismatches 432; Indels 24; Gaps 4;	Qy	1082	CACAGCCCTAGCAAAAGACAGGTCTACACAGATGTTGGCTGATGAGATGACCATCA 1141
Qy	5	Db	1804	AACCTGCTTGGCTTAAAGACAAAGGCNTACGACAGATTCATCTTGTATGACATGACAGTTCA 1863
Db	745	Qy	1142	CAAGGCTTTGACGCTTGGCAAGATGCCAATACACCTTTTGGATTTCTTTAACGGCCACAG 1201
Qy	65	Db	1864	CMAGGCTTACACTTGGGCAAGATGCGAGCCGCTTATGGAATCTTCAACGGGACAG 1923
Db	805	Qy	1202	ATCCAGATGTGCTCCGCTCTGATCCTTTGCTGAAGGTGATGAGAGCACTGGCTAATCC 1261
Qy	125	Db	1924	ATGTCACATGTGCTTGGCTCAGACTCTCTTGTGAAAGTATGATGAGCGGTGGCGAATCC 1983
Db	865	Qy	1262	TGGGCTGGCGCGGTGTTTCATTTGGAAGCTGGGAGCAACAGTGTGAGGCGCATATC 1321
Qy	185	Db	1984	AGGTGTAAGAGGCTGGTGATAGTGAAGCAGGAGCAACAGTGTGGAAGGTATCATTTTC 2043
Db	925	Qy	1322	ACTAAGTGACATTTTCAAGTTCTTGTGAGCTTGTGA 1358
Qy	245	Db	2044	ACTGAGGATGTTTCCAAATTCCTGCTCGTCTTTGA 2080
Db	979	Qy		
Qy	305	Db		
Db	1024	Qy		
Qy	365	Db		

RESULT 8
AF439826
LOCUS
DEFINITION
Arabidopsis thaliana At:g09020/F7G1.9.11 mRNA, complete cds.
ACCESSION
AF439826

VERSION	AF439826.1	GI:16612254
KEYWORDS	PLZ CDNA	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Shinn, P., Chen, H., Cheuk, R., Kim, C.-J., Koesema, E., Meyers, M. C., Banth, J., Bowser, L., Carninci, P., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jiang, P. X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.	
AUTHORS	Arabidopsis cdna clones	
TITLE	Unpublished	
REFERENCE	2 (bases 1 to 2301)	
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.-J., Koesema, E., Meyers, M. C., Banth, J., Bowser, L., Carninci, P., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jiang, P. X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-OCT-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA' : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.	
FEATURES	<p>The Salk, Stanford, FGENC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C.-J., Koesema, E., Meyers, M. C., Banth, J., Bowser, L., Chang, E., Dale, J. M., Goldsmith, A. D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R. W., Theologis, A., and Ecker, J. R.</p> <p>Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.</p> <p>Location/Qualifiers</p> <p>1..2301</p> <p>/organism="Arabidopsis thaliana"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:3702"</p> <p>/chromosome="1"</p> <p>/clone="RAFL09-86-B11 (R14970)"</p> <p>/ecotype="Columbia"</p> <p>1..603</p> <p>604..2067</p> <p>/note="putative activator subunit of SNF1-related protein kinase SNF4"</p> <p>/codon_start=1</p> <p>/product="At1g09020/F7G19.11"</p> <p>/protein_id="AAL27498.1"</p> <p>/db_xref="GI:16612255"</p> <p>/translation="MFGSLDSSRGNSAAGSGLLPTFRVWVYGGRRVFLSGFTRWIT EHVPMPLGEGCTVFQVINCILNLTGPHQXKFVDGWRHDEHOPFVSGNGVYVITPIT GPDVMPAGFSPETLGRSNMVDVFLRTADPSQEAAPVPMRSGVLDLSRRHRSVLLSTR TAYELLPSGKVIADVNLVPRKAFHLYEYQIPLAPWDFGKGFVGVLPDILFILLI LRELTHGNSVTESELEYTHIAAKEGKAHISROYDGSGRPYPRFLVQVGPDLKDV ALKILONKVAAPVLYSSIDGSPYQLHLASLILKICRYFRHSSSLPILQOPI CSIPLTWVPRIGSSSSKPLATLRPHASIGSALALLVQAEVSSIIPWVDNDLSLIDIS</p>	
5'UTR	<p>misc_difference 1267</p> <p>3'UTR</p> <p>ORIGIN</p> <p>Query Match 36.7%; Score 577.8; DB 8; Length 2301;</p> <p>Best Local Similarity 66.4%; Pred. No. 1.1e-148;</p> <p>Matches 901; Conservative 0; Mismatches 432; Indels 24; Gaps 4;</p>	
CDS	<p>1267</p> <p>/note="compared to genome"</p> <p>/replace="g"</p> <p>2068..2301</p> <p>5 GAGGACACATTCGGGATCTCTCTATCGAGGCTGCCACCTGATTTTCAAGGCTATTG 64</p> <p>732 GACAGAGATGTCGAATGTCACCGCTCGAGGCTGCCCTCTCTCTCTCTCTCTCTCT 791</p> <p>65 CAGTCTGCT 124</p> <p>792 CAACCTTGAGCCGAGATATCATCATGATTAAGTTTTTTTGTGATGGGAATGGCGCAGA 851</p> <p>125 TGAGGCTCAACCTACCATATCTGGGAGTTTGGCATAGTTAAACACACTTTACTTTGACAAG 184</p> <p>852 TGAGCACCACCACTTTGTAGCGGGAATGGTGGAGTAGTGAATACAAATTTTATAACTGG 911</p> <p>185 GGAATATAACCAATAAACAACCTTTATCAAGTCCCAAGCACACTGGAAGCAGGATGAACAT 244</p> <p>912 GCCAGATA---TGCT 965</p> <p>245 GGATGTGATTAATGAAATTTTCAACGCTACGGTTACGTTGTGAGANGCACCCTTTCAGA 304</p> <p>966 GGATGTGATGATGTTCT 1010</p> <p>305 AGGTACTCTGAGAGTTTCAGAGGCTGCATACAAATATCTAGGTCTCTCTCTCTCTCTGAATA 364</p> <p>1011 AGCTGTTCTCTAGAGTGTGGGGGTTGATTTGGAGCTTTCTCTCTCACCAGATATCGGTTT 1070</p> <p>365 TCTGAATTTGCATACATCATGATTTACTCCAGATTTCTGCCAGATTTCTGCCAAGGTTTATTC 424</p> <p>1071 ATTGTCAACCGCACTGCATATGAGTCTCTCCAGATTCGGCAAGTTTATTCGATTGGA 1130</p> <p>425 CATTAATTTACCTGTGAAGCAATCATTCATTTCCATGAACAGGGGATTCCTGTAGC 484</p> <p>1131 TGTAAATTTGCCAGTGAAGCAAGCATTCATATCTCTATGACGAGGGAATCCCTTTGGC 1190</p> <p>485 TCCTCTCTGGAGCTCATTCAGAGGCTCAATTTGTTGGTCTCTCTTACCCATTCGATTTTCTAT 544</p> <p>1191 TCCTCTTTGGGACTTTGGAAAAGGCCAATTTGTTGGAGTTCTTGTCTCCTACATGACTTCTAT 1250</p> <p>545 ACTCATATTGGGAGCTAGAAAACCTCATGGCTCGAACTTGACAGAGAGCAGCTTGAAC 604</p> <p>1251 TCTAATCTGAGAGAGCTTGGAACTCATGTATCAAACTTGACAGAGAGAGCTTGGAGAC 1310</p> <p>605 ACACACTATATCTGATGGAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATAGTCA 664</p> <p>1311 GCACACAATAGCAGCTCGAAAAGAGGGGAGGCTCATATTAGCCGGCAATATGATGGAG 1370</p> <p>665 GTGGGCGCGCAAC---AGCATCTAGTGCATGCCACCCCTTATGAGTCTCTTGGAGGACAT 721</p> <p>1371 TGGGAGACCATATCTCTAGGCACTTGTTCAGTTGGAACCTATGATGATTAATCTGAAGACCT 1430</p> <p>722 TGCACTAAAGCTTTTSCAAAATGCGCAATTTCTACAGTGGCAGTTATTTATTCATCATCATC 781</p> <p>1431 TGCCCTGAAAATTTTGGAAAACAAGGTGGCAGCTGTCGGTTATATATTTCTTCTTGA 1490</p> <p>782 AGATGATCATTCGCGAGTTATTCATCTCTGATCATCTCTCTGATTTCTGGAATTTGCAATGAT 841</p> <p>1491 GGATGGTTTATATTCGCGAGTTTACTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1550</p> <p>842 TTGTAGATCTTCAAAAACCTCAACTGGTAAATTTGGCTATTCTGAAACCAACCAAGTGTGCTC 901</p> <p>1551 ATGCAGATCTTTAGACATTCGTCTAGCTCTTTTACCTATCTCTCTCTCTCTCTCTCTCTCT 1610</p> <p>902 CATTCGCTGGGTTCTCTGGGTTCCGAAATTCGGTATTCGAAACAGTGTGCTTCAATTCGCTAT 961</p>	

1611	AA	TCCCTCGGTACG	TGGGTCCCTAGAA	TCGGAGAA	TCAAGTAGCAAA	CCTCTCGCTAC	16170
962	GT	TGCGACCTAA	TGCGCTCACT	TAGCTCTG	CCCTTAACATG	TGCGTTCAAGCTCGAGTAAG	1021
1671	ATT	GAGACCTTCAC	GCCTCTTTTGGG	TCTCGCGTCG	CGAATTAAT	TAGTTCAAGCTCGAATCG	1730
1022	CTCAAT	ACCAAATGTGGAT	CAACAAGACTCCCT	CGCTTTCACACT	TACTCTAGAAAGTGACAT		1081
1731	TTCAAT	TCCGTAGTGGAT	GACACAGCTCGCT	TATTGACAT	TACTCTCGAAGTGATAT		1790
1082	CACAGC	CCCTAGCAAAAGACA	AGGTCTTACACACAT	TGTTGGCTCGAT	GAGATGCCAATCA		1141
1791	AAC	TGCTCTGGCTAAAGACA	AGGCATACGACACAGAT	TTCATCTTGAT	GACATGACAGTTCA		1850
1142	CGAGG	CTTTGCAGCTTGGACA	AGATGCCAATACACCT	TTTGGATCTTTTAA	CGGCCAGAG		1201
1851	CCAGG	GC3CTACAGTTGGGGCAAGAT	CGAGCCGCTTAT	TGGAATCTTCAAC	CGGGCAGAG		1910
1202	ATGCGA	GTGCTCGCGTCTGAT	CTTTGCTGAAGCTGAT	GAGCGATGGCTGAATCC			1361
1911	ATGT	CATATGCTTGGCGCTCAG	ACTCTTTGTGAAGTGA	TGAGGCGGTTGGCGAATCC			1970
1262	TGGGGT	TCGGCGGGGTGTTCAT	TGTGGAAAGCTGGGAGCAAA	CGTGTGGAGGGCATCATC			1321
1971	AGGTGA	TAAAGAGGCTGGTGAT	TGGGAAGCAGGAGCAAA	CGTGTGAAGGATATTC			2030
1322	ACTAAG	TGACATTTTCAAGTCT	TGCTGAGCTTGTA				1358
2031	ACT	GAGCGATTTTCCAAATCT	CTGCTCGGTCTTTGA				2067

RESULT	9
LOCUS	AY247268
DEFINITION	Medicago truncatula AKIN betagamma mRNA, complete cds.
ACCESSION	AY247268
VERSION	AY247268.1 GI:29887974
KEYWORDS	.
SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Medicago truncatula
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE	1 (bases 1 to 1887) Buitink,J., Thomas,M., Gissot,L. and Leprince,O. Differential expression patterns of beta and gamma subunit isoforms suggest the involvement of different ScRKI complexes in desiccation tolerance, osmotic stress and starvation in germinating seeds of Medicago truncatula Unpublished
JOURNAL	2 (bases 1 to 1887) Buitink,J. and Leprince,O. Direct Submission
REFERENCE	Submitted (03-MAR-2003) UMR 1191 Molecular Seed Physiology, INRA, 16 Bd Lavoisier, Angers 49045, France Location/Qualifiers
AUTHORS	
TITLE	
JOURNAL	
FEATURES	

```

1. .1887
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="Paraggio"
/db_xref="taxon:3880"
/dev_stage="embryo"
65. .1522
/codon_start=1
/product="AKIN betagamma"
/protein_id="AA061673.1"
/db_xref="GI:29887975"
/translation="MPSLRDSVGGSSLSSSSGPLIPKRFWYPYGGTRVLLIGSF
TRSEHLPMSMEGCGPSVFOVLSLMEGHQFNVDGQWRKYDEQQPFVNGNTGIVNT
IVLRPDDIPLALISAETSSRHEVNDVNGHAEANPRMSPDLEVRRI1SKFLSE
HTAVDILPESGKIALVDNVLPAQAFHLYVEODVSMAPLWDFCKSFQVGLSAMDFLL

```

[illegible]

Db	1073	CCTTGGGTACATGGGTGCTAAACGTTGGGATCCAAATGGACAGCCACTGTAAGGCTA	1132	Db	132	CGTTTCGGCAGCTGTTGCACTCTTGGCATCCCTTTGAGGAATTTTGAATGTATCTGTAGAT	191
Qy	966	CGACCTAATGCCCTCACTTAGCTCTGCGCTTAAACATCTTGGTTCGAAGCTGGAGTAAGCTCA	1025	Qy	850	ACTTCAAAACTACACTGTTGTAATTTGGCTATTCTGAACCAACAGTGTGCTCCATTCCGC	909
Db	1133	AGCCAAATGCTTCTCTTGGTGTATGCCCTCTCGATGTTTGTTCGAAGCTAAAGTCAGCTCA	1192	Db	192	ACTTCAAGAACTCCACTGGTAGTTTGGCGATTCTTAAACCAACAGTAGTGTCTCAATTCCGC	251
Qy	1026	ATACCAATTTGGTATGACCAAGCACTCCCTGCTTGACACTTACTCTAGAACTGACATCACA	1085	Qy	910	TGGGTTCTCGGGTTCGGAATTTGGTGTATCTGAACAGCTCGTCCATTGGCTATGTTGGCAC	969
Db	1193	ATACCAATTTGGTATGAAATGATTCATTGCTTGACATATATTCAAGAGTGTATTAAT	1252	Db	252	TGGGTACTCGGGTTCGGAATTTGGTGTATCTGAACCAATGCTCATCCATTGGCTATGTTGGGC	311
Qy	1086	GCCTTAGCAAGACAGCTCTACACATGTTCCGGCTGGATGAGTACCATTCACACAG	1145	Qy	970	CTAATGCTCCTTACCTTAGCTCTGCCCTTAACTTGGTTCGAAGCTGGAGTACGATATAC	1029
Db	1253	GCATTTGCTAAGATAAGCATATGCTCGATATCCCTGGATGAACCTAATATTCATCAG	1312	Db	312	CTAATACATCTCTTAGCTCTGCCCTTAACTTGGTTCGAAGCTGGAGTACGATATAC	371
Qy	1146	GCTTTGCACTTGGCAAGATGCCAATACCTTTTGGATCTTTTAAAGCCGACAGATGC	1205	Qy	1030	CAATTTGGATGACAGCACTCCCTGCTTGACACTTACTCTAGAAGTGCATACACAGCC	1089
Db	1313	GCATTAATTTGGCAAGATGCCAATATCTCTTATGCTTAAACATGGCCACAGATGT	1372	Db	372	CCATTTGCTGATGAACGACTCGCTGATCGACATATCTCCAGAACTGACATACAGCTC	431
Qy	1206	CAGATGTCCTCGGCTGATCTCTTTGCTGAAGTGTAGAGCGACTGGCTAATCTCGG	1265	Qy	1090	TAGCAAAAGACCAAGCTCTACACATGTTCCGCTGGATGAGATGACCAATTCACAGGCTT	1149
Db	1373	CATATGTTTGGCTGATCTGATTCATTCACAAAGTGTAGAGCGGTTGGCTAAACCTGG	1432	Db	432	TAGCAAAAGACCAAGCTCTACACATGTTCCGCTGGATGAGATGACCAATTCACAGGCTT	491
Qy	1266	GTGCGGGGCTGTTCTAGTGGAGCTGGAGCAACGTTGGAGGGCATCATATCACTA	1325	Qy	1150	TGCAGCTTGGCAAGATGCCAATACACTTTTGGATCTTTTAAAGCCGACAGATGCCAGA	1209
Db	1433	GTTAGGAGACTGTTGATTTGGAACTGGTAGCAAGCTGTGGAGGGATTTATTCATTA	1492	Db	492	TGCAGCTTGGCAAGATGCCAATACACTTTTGGATCTTTTAAAGCCGACAGATGCCAGA	551
Qy	1326	AGTGACATTTTCAAGTTCTTCTGCTGAGCTTGTGAGGAGAGA	1365	Qy	1210	TGTCCTCGGCTGATCTCTTGGTGTAGAGTGTAGAGCGACTGCTAATCTCGGCTGC	1269
Db	1493	AGTGATCTTTTCAAGTTCTTCTGCTGAGCTTGTGAGGAGAGA	1532	Db	552	TGTCCTCGGCTGATCTCTTGGTGTAGAGTGTAGAGCGACTGCTAATCTCGGCTGC	611
RESULT 10				Qy	1270	GGCGGGTTCATTTGTTGAAGCTGGAGCAACGTTGGAGGGCATCTATCATCACTAATGTG	1329
BT009623				Db	612	GTCGCTGTTTCATCTGTCGAGGCTGTCGACAGCGAGTGGAAAGCGCTAATCTGCTGAGCG	671
LOCUS				Qy	1330	ACATTTTCAAGTTCTTCTGCTGAGCTTGTGA	1358
DEFINITION				Db	672	ACATTTTCAAGTTCTTCTGCTGAGCTAGCGA	700
ACCESSION				RESULT 11			
VERSION				AK1211122			
KEYWORDS				LOCUS			
SOURCE				DEFINITION			
ORGANISM				ACCESSION			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				VERSION			
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				KEYWORDS			
1 (bases 1 to 982)				SOURCE			
Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,				ORGANISM			
Caraher,N.R., Hanafey,M.K. and Hainey,C.F.							
Direct Submission				REFERENCE			
Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and				AUTHORS			
Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,							
USA							
FEATURES							
source							
1..982							
/organism="Triticum aestivum"							
/mol_type="mRNA"							
/db_xref="taxon:4565"							
/clone="wreln.pk0143.e2.fis"							
ORIGIN							
Query Match							
Best Local Similarity							
Matches 587; Conservative							
0; Mismatches 102; Indels							
0; Gaps							
0;							
Qy	670	GACCGAACAGCATCTAGTGTGCGACCCCTTATGAGTCTTGGAGGACATGCACTAA	729				
Db	12	GATCAATCAGCATCTAGTGTGCGACCCCTTATGAGTCTTGGAGGATTTGCCATGA	71				
Qy	730	AGCTTTTGAATGCAATCTTCTACAGTCCAGTATTTTATTCATCATCATCATGAGTAT	789				
Db	72	AAATCTCGAACTGCAATCTTCTACAGTCCCAATCATCTATTCATCATCATCATGAGTAT	131				
Qy	790	CATTCCCGCAGTATTGCACTCTTGCATCACTTTCTGGAATTTTGAATGTATTGTAGAT	849				

TITLE
JOURNAL
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)

Query Match	15.3%	Score 241.6	DB 8	Length 1923
Best Local Similarity	54.0%	Pred. No. 1.7e-55		
Matches 552	Conservative	0	Mismatches 449	Indels 21
22752273				
12689764				
2				
ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imoto, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakabana, Y., Nakamura, M., Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.				
3	(bases 1 to 1923)			
Direct Submission				
Submitted (31-JAN-2003)	Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)			
35	This clone is one of the 32K full-length cDNA clones from japonica rice.			
URL:	http://cdna01.dna.affrc.go.jp/cDNA/NTAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, M., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M. and Nakahama, Y.			
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.				
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hayashizaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.				
Location/Qualifiers				
1	1923			
/organism="Oryza sativa (japonica cultivar-group)"				
/mol_type="mRNA"				
/cultivar="Nipponbare"				
/db_xref="taxon:39947"				
/clone="J023074117"				
FEATURES				
source				
ORIGIN				

```

Db      1825 1826

RESULT 12
AC128647 92392 bp DNA linear PLN 04-SEP-2003
LOCUS Oryza sativa chromosome 3 BAC clone OSUNB0062G19, complete
DEFINITION
AC128647
VERSION AC128647.8 GI:34447243
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 92392)
AUTHORS Buell,R., Hsiao,J., Zismann,V., Moffat,K.M., Hill,J.,
Gansberger,K., Burgess,S., Jarrahi,B., Shwartsbeyn,M., Brenner,M.,
Ciecko,A., Pai,S., Vanaken,S., Hansen,C., Utterbach,T.,
Feldblyum,T., Khalak,H.G., Yuan,Q., Quackenbush,J., White,O.,
Salzberg,S. and Fraser,C.
TITLE Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic
sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 92392)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 92392)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 92392)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
COMMENT On Sep 4, 2003 this sequence version replaced gi:28827864.
FEATURES
    source
        1..92392
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="genomic DNA"
            /cultivar="Nipponbare"
            /sub_species="japonica"
            /db_xref="taxon:39947"
            /chromosome="3"
            /clones="OSUNB0062G19"

ORIGIN
Query Match 8.6%; Score 135; DB 8; Length 92392;
Best Local Similarity 64.4%; Pred. No. 5.5e-26;
Matches 266; Conservative 0; Mismatches 65; Indels 82; Gaps 1;

Qy      683 TCTAGTCATGCCACCCCTTATAGTCTCTTACGAGCAATTCGAGGCAATTCGAGCTTTTGCATAA 742
Db      79838 TATTGTTAGGCCACCCCTTATGAATCTCTGAGGGAATTCGTATGAGATCTGCAAA 79897

Qy      743 TGGCATTTCTACATGCGCAGTTATTTATTCATCATCATCATCATTCATTCGCCAGTT 802
Db      79898 CGGTGTTCTACCGTTCCAAATATGTTTCTCTCATCCAGATGCTCATATCCAAAT 79957

Qy      803 ATTGCATCTTGCATCAGCTCTTCTGGAATTTGAAAT----- 837
Db      79958 GTTCACTCTTCTCCCTTCTGGAATTTTGAATGTAAGTGTTCATGATTTCGATG 80017

Qy      838 -----GTA 840
Db      80018 AAAAGTGTGTATACAAACATATTATCATGATAGCTTAAACAATTTCAATTTTCAGGTA 80077

```

Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by repeatmasker (Arian Smith, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

This BAC overlaps with rice BAC OSJNBb0062G19 (AC128647) and OSJNBa0033P04 (AC092263).

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital Status</i>	3. <i>Marital Status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Occupation</i>	5. <i>Occupation</i>
6. <i>Income</i>	6. <i>Income</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political Affiliation</i>	8. <i>Political Affiliation</i>
9. <i>Health Status</i>	9. <i>Health Status</i>
10. <i>Travel History</i>	10. <i>Travel History</i>
11. <i>Employment Status</i>	11. <i>Employment Status</i>
12. <i>Family Size</i>	12. <i>Family Size</i>
13. <i>Home Ownership</i>	13. <i>Home Ownership</i>
14. <i>Vehicle Ownership</i>	14. <i>Vehicle Ownership</i>
15. <i>Insurance Status</i>	15. <i>Insurance Status</i>
16. <i>Subscription Services</i>	16. <i>Subscription Services</i>
17. <i>Charitable Contributions</i>	17. <i>Charitable Contributions</i>
18. <i>Volunteer Work</i>	18. <i>Volunteer Work</i>
19. <i>Real Estate Transactions</i>	19. <i>Real Estate Transactions</i>
20. <i>Banking History</i>	20. <i>Banking History</i>
21. <i>Investment Portfolio</i>	21. <i>Investment Portfolio</i>
22. <i>Medical History</i>	22. <i>Medical History</i>
23. <i>Travel Itinerary</i>	23. <i>Travel Itinerary</i>
24. <i>Employment History</i>	24. <i>Employment History</i>
25. <i>Family Medical History</i>	25. <i>Family Medical History</i>
26. <i>Home Renovation Projects</i>	26. <i>Home Renovation Projects</i>
27. <i>Vehicle Maintenance Records</i>	27. <i>Vehicle Maintenance Records</i>
28. <i>Insurance Claims History</i>	28. <i>Insurance Claims History</i>
29. <i>Subscription Cancellations</i>	29. <i>Subscription Cancellations</i>
30. <i>Charitable Donations</i>	30. <i>Charitable Donations</i>
31. <i>Volunteer Hours</i>	31. <i>Volunteer Hours</i>
32. <i>Real Estate Listings</i>	32. <i>Real Estate Listings</i>
33. <i>Banking Statements</i>	33. <i>Banking Statements</i>
34. <i>Investment Performance</i>	34. <i>Investment Performance</i>
35. <i>Medical Records</i>	35. <i>Medical Records</i>
36. <i>Travel Itineraries</i>	36. <i>Travel Itineraries</i>
37. <i>Employment History</i>	37. <i>Employment History</i>
38. <i>Family Medical History</i>	38. <i>Family Medical History</i>
39. <i>Home Renovation Projects</i>	39. <i>Home Renovation Projects</i>
40. <i>Vehicle Maintenance Records</i>	40. <i>Vehicle Maintenance Records</i>
41. <i>Insurance Claims History</i>	41. <i>Insurance Claims History</i>
42. <i>Subscription Cancellations</i>	42. <i>Subscription Cancellations</i>
43. <i>Charitable Donations</i>	43. <i>Charitable Donations</i>
44. <i>Volunteer Hours</i>	44. <i>Volunteer Hours</i>
45. <i>Real Estate Listings</i>	45. <i>Real Estate Listings</i>
46. <i>Banking Statements</i>	46. <i>Banking Statements</i>
47. <i>Investment Performance</i>	47. <i>Investment Performance</i>
48. <i>Medical Records</i>	48. <i>Medical Records</i>
49. <i>Travel Itineraries</i>	49. <i>Travel Itineraries</i>
50. <i>Employment History</i>	50. <i>Employment History</i>
51. <i>Family Medical History</i>	51. <i>Family Medical History</i>
52. <i>Home Renovation Projects</i>	52. <i>Home Renovation Projects</i>
53. <i>Vehicle Maintenance Records</i>	53. <i>Vehicle Maintenance Records</i>
54. <i>Insurance Claims History</i>	54. <i>Insurance Claims History</i>
55. <i>Subscription Cancellations</i>	55. <i>Subscription Cancellations</i>
56. <i>Charitable Donations</i>	56. <i>Charitable Donations</i>
57. <i>Volunteer Hours</i>	57. <i>Volunteer Hours</i>
58. <i>Real Estate Listings</i>	58. <i>Real Estate Listings</i>
59. <i>Banking Statements</i>	59. <i>Banking Statements</i>
60. <i>Investment Performance</i>	60. <i>Investment Performance</i>
61. <i>Medical Records</i>	61. <i>Medical Records</i>
62. <i>Travel Itineraries</i>	62. <i>Travel Itineraries</i>
63. <i>Employment History</i>	63. <i>Employment History</i>
64. <i>Family Medical History</i>	64. <i>Family Medical History</i>
65. <i>Home Renovation Projects</i>	65. <i>Home Renovation Projects</i>
66. <i>Vehicle Maintenance Records</i>	66. <i>Vehicle Maintenance Records</i>
67. <i>Insurance Claims History</i>	67. <i>Insurance Claims History</i>
68. <i>Subscription Cancellations</i>	68. <i>Subscription Cancellations</i>
69. <i>Charitable Donations</i>	69. <i>Charitable Donations</i>
70. <i>Volunteer Hours</i>	70. <i>Volunteer Hours</i>
71. <i>Real Estate Listings</i>	71. <i>Real Estate Listings</i>
72. <i>Banking Statements</i>	72. <i>Banking Statements</i>
73. <i>Investment Performance</i>	73. <i>Investment Performance</i>
74. <i>Medical Records</i>	74. <i>Medical Records</i>
75. <i>Travel Itineraries</i>	75. <i>Travel Itineraries</i>
76. <i>Employment History</i>	76. <i>Employment History</i>
77. <i>Family Medical History</i>	77. <i>Family Medical History</i>
78. <i>Home Renovation Projects</i>	78. <i>Home Renovation Projects</i>
79. <i>Vehicle Maintenance Records</i>	79. <i>Vehicle Maintenance Records</i>
80. <i>Insurance Claims History</i>	80. <i>Insurance Claims History</i>
81. <i>Subscription Cancellations</i>	81. <i>Subscription Cancellations</i>
82. <i>Charitable Donations</i>	82. <i>Charitable Donations</i>
83. <i>Volunteer Hours</i>	83. <i>Volunteer Hours</i>
84. <i>Real Estate Listings</i>	84. <i>Real Estate Listings</i>
85. <i>Banking Statements</i>	85. <i>Banking Statements</i>
86. <i>Investment Performance</i>	86. <i>Investment Performance</i>
87. <i>Medical Records</i>	87. <i>Medical Records</i>
88. <i>Travel Itineraries</i>	88. <i>Travel Itineraries</i>
89. <i>Employment History</i>	89. <i>Employment History</i>
90. <i>Family Medical History</i>	90. <i>Family Medical History</i>
91. <i>Home Renovation Projects</i>	91. <i>Home Renovation Projects</i>
92. <i>Vehicle Maintenance Records</i>	92. <i>Vehicle Maintenance Records</i>
93. <i>Insurance Claims History</i>	93. <i>Insurance Claims History</i>
94. <i>Subscription Cancellations</i>	94. <i>Subscription Cancellations</i>
95. <i>Charitable Donations</i>	95. <i>Charitable Donations</i>
96. <i>Volunteer Hours</i>	96. <i>Volunteer Hours</i>
97. <i>Real Estate Listings</i>	97. <i>Real Estate Listings</i>
98. <i>Banking Statements</i>	98. <i>Banking Statements</i>
99. <i>Investment Performance</i>	99. <i>Investment Performance</i>
100. <i>Medical Records</i> </	

source	repeat_region	gene	mRNA	CDS
1..145115				
/organism="Oryza sativa (japonica cultivar-group)"				
/mol_type="genomic DNA"				
/cultivar="Nipponbare"				
/db_xref="taxon:39947"				
/chromosome="3"				
/map="near L375"				
/clone="OSUNBa0059G06"				
/note="japonica cultivar-group"				
1271..1331				
/rpt_family="AT_rich"				
2342..3105				
/gene="OSUNBa0059G06.1"				
/note="predicted by genemarkEMM"				
join(2342..2619,2999..>3105)				
/gene="OSUNBa0059G06.1"				
join(<2342..2619,2999..3105)				
/gene="OSUNBa0059G06.1"				
/codon_start=2				
/product="hypothetical protein, 5'-partial"				
/protein_id="AAP21395.1"				
/db_xref="GI:30102982"				
/translation="EQMREAAGGGGHCRCRWARERRCOLREGVEAAAGGGRARRGEGGCRHRTAASSSHIVAREIQPELSNLLFLSSWWSASAVDPQASGAVNDSERLHRPAASEAATARGASASTVGR"				
2353..2479				
/rpt_family="G-rich"				
3392..3467				
/rpt_family="(CCG)n"				
5230..5343				
/rpt_family="(TATAA)n"				
5868..8753				
/gene="OSUNBa0059G06.2"				
/note="similar to family II lipase EXL3 GB:AAK30018 GI:15054386 (Arabidopsis thaliana); contains Pfam profile PF00657 (GDSL like Lipase Acylhydrolase)"				
join(<5868..6150,7939..>8753)				
/gene="OSUNBa0059G06.2"				
join(5868..6150,7939..8753)				
/gene="OSUNBa0059G06.2"				
/codon_start=1				
/product="putative lipase/acylhydrolase"				
/protein_id="AAP21383.1"				
/db_xref="GI:30102970"				
/translation="MTMKQKQNMKGLLVAVAVAPAGSPASSGSGAPMPVAVLVEGSDVDRGNMNLKTIQIKANHAPVGMDFANSPKGRYSNGLIPDTDFIVQGLNKGIMPPYGLVELSPDLKTVGSFASGATGYDELTPVIVSVITLDQQLIEFHEHRYKRLVDGVGVEETARIIGALFVWCAGTDITANTYITPFRSVETDIPSVLDLVSQAARKLDGVAAALGARRGVFLPIQLQELVNGSQRTLGCGHPRECKRYAAKLFNSMEEVIAKTNITATRMVYDIITLIQELVENGDKYGTFTTTRGCGGTGTIEVTGLCDARFVDICDNVSNHVFPSYHPTQRAYKIIVDYIPDNLQFLA"				
6410..6437				
/rpt_family="AT_rich"				
7781..7804				
/rpt_family="(CGAT)n"				
12122..13806				
/gene="OSUNBa0059G06.3"				
/note="predicted by genesh"				
join(<12122..12362,12851..13020,13702..>13806)				
/gene="OSUNBa0059G06.3"				
join(12122..12362,12851..13020,13702..13806)				
/gene="OSUNBa0059G06.3"				
repeat_region				
repeat_region				
gene				
mRNA				
CDS				

```

/codon_start=1
/product="hypothetical protein"
/protein_id="AAP21385.1"
/db_xref="GI:30102972"
/translation="MRVFKGKVCAGPFGKESVSGAGNGGAPRRAGDERPPGADGNG
SEAPIDGAGKMFGEVGLSSSTVFARVVDELGPDDARNMAMCMCAAKAVDTAEF
SSRCGLARSQESWQPCSELAARRQSMTVAEVELDSGNRESRDGFKVDSDIRGIG
TSINWRFGKGD"
repeat_region
13321..13364
/rpt_family="GA-rich"
repeat_region
15008..15049
/rpt_family="(TC)n"
repeat_region
16325..16352
/rpt_family="(TTCC)n"
misc_feature
16964..17005
/note="Overlap BAC OSJNBa0033P04 has 49 TA repeats in this
region.*"
repeat_region
16964..17005
/rpt_family="(TA)n"
repeat_region
17177..17203
/rpt_family="AT-rich"
gene
17384..18823
/gene="OSJNBa0059G06.4"
/note="contains Pfam profile PF01554 [Mate]"
mRNA
<17384..>18823
/gene="OSJNBa0059G06.4"
CDS
17384..18823
/gene="OSJNBa0059G06.4"
/codon_start=1
/product="putative Mate efflux family protein"
/protein_id="AAP21388.1"
/db_xref="GI:30102975"
/translation="MTPTPPSPHPERTWAESVASPEPRAQGIAPFLIAMLTFEKL
AVTAPFLGRGLQALQALGTSFANVTGAVTGLICAMDPICQGHGASNGKLKX
TLVMATILLLGASITLFIASALHVALPLTMMARTGIOGVATAVWSDLAVMILA
PLKSYLSAGQVTLPTLFASALALHPLTMMARTGIOGVATAVWSDLAVMILA
GVVLVSERRKAGGCGGQWGEVGRVLRLLAVPSCLNTCLEWCCYEILVLTGRLP
DARRTVAVMAVTNFDYLLFAGMLSLVSASVAVSBLGAGEAWARRAGMVSIVGGA
VCGVGGVGAWAARRAGWSIYSSDAGVREGVGRAMEVMVLEVVNPLNVCGGIVRG
APVAGVMYAVAGFYVIALPGLVALFAKRLGQLGLGLVGNASLAVLTLFIARM
DWPAPAKARTATRTATVAQFHQHDEVVQP"
repeat_region
18325..18515
/rpt_family="G-rich"
gene
20002..21960
/gene="OSJNBa0059G06.5"
/note="EST CA759271,BC040176 from this gene"
mRNA
Join<20002..21123,20424..20298,20375..20500,20579..20654,
20811..20903,21132..21245,21366..21960>
/gene="OSJNBa0059G06.5"
CDS
Join(20002..20123,20424..20298,20375..20500,20579..20654,
20811..20903,21132..21245,21366..21419)
/gene="OSJNBa0059G06.5"
/codon_start=1
/product="unknown protein"
/protein_id="AAP21391.1"
/db_xref="GI:30102978"
/translation="MAPPPPCCVLWVRCCSGLSPFPSPPPPPPSNNLRIRAREFV
LSSSLATIAIPLHLSGTPRYGVQKSPSLCAPTNVCSTCEDITDSIHVAPPW
NYPKQGRRAKPTIHEALNQLQVYQCKPDNFTPELVKTDDYVEVESPIEGFV
DOVEWFPFGKNSIVYRSASRSRGFDENAKKRVKALRLALENKGWASESTI"
repeat_region
20063..20084
/rpt_family="(CCG)n"
repeat_region
20673..20703
/rpt_family="(TTTC)n"
repeat_region
21163..22193
/rpt_family="AT-rich"
gene
24255..25676
/gene="OSJNBa0059G06.6"
/note="predicted by fgenesh"

```

Query Match 8.6%; Score 135; DB 8; Length 145115;
Best Local Similarity 64.4%; Pred. No. 5.5e-26;
Matches 266; Conservative 0; Mismatches 65; Indels 82; Gaps 1;

Query Match 8.6%; Score 135; DB 8; Length 145115;
Best Local Similarity 64.4%; Pred. No. 5.5e-26;
Matches 266; Conservative 0; Mismatches 65; Indels 82; Gaps 1;

QY	683	TCTAGTGCATGCCACCCCTTATGAGTCCCTTGAGGGACATTCGACGTAAGACTTTTGC AAAA	742
Db	135226	TATGTGTTTAGGCCACCCCTTATGAATCCCTGAGGGAAATTCCTATGAAGATACCTGCAAAA	135167
QY	743	TGCAATTTCTACAGTSCAGTTATTTATCATCATCATCAGACGATCATTCCTCCGCGATT	802
Db	135166	CGGTGTTTCTACCGTTCCAAATATGTTTCTCATCACAGATGGCTCATATCCACAAT	135107
QY	803	ATTGCATCTTGCATCAGCTTTCTGGAAATTTGAAAT	837
Db	135106	GTTCATCTTGTCCCTTTCTGGAAATTTGAAATGTAAGTGTCTCTATGATTATTCGATG	135047
QY	838	-----GTA	840
Db	135046	AAAAAGTGTGCTATACACAAACATATTACATGATGCTTAAACAAATTCATTTTCAGGTA	134987
QY	841	TTTGTAGATACTTCAAAACCTCAACTGCTTAATTTGCTTATTTCTGAACCAACAGGTGCT	900
Db	134986	TTTGTAGATATTTAAAATTTCTCAAGTAAATTTACCTATTTTGAGCCAACTGATGCA	134927
QY	901	CAATCCGCTGGGTCTCTGGTTCGGAATTTGGTGATCTCGAACAGTCGTCTCATTTGGCTA	960
Db	134926	CAATTCCTCTGGTACCTGGTTCGAAATTTGGTGATCTCAATGGCGCTCATTTGGCTA	134867
QY	961	TGTTGGACCTCAATGCTCTAGTCTGCGCTTAACATGTTGTTCAAGCT	1013
Db	134866	TGTTGGGCTTACACATCTCTTAGCTGCGCTGAATTTGCTGGTTCAAGT	134814
RESULT 14			
AC146552/c			
LOCUS	AC146552	107857 bp	DNA linear HTG 30-AUG-2003
DEFINITION	Medicago truncatula clone mth2-9d1, WORKING DRAFT SEQUENCE, 9		
ACCESSION	AC146552		
VERSION	AC146552.1	GI:34365866	
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT.		
SOURCE	Medicago truncatula (barrel medic)		
ORGANISM	Medicago truncatula		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;		
	Medicago.		
REFERENCE	1 (bases 1 to 107857)		
AUTHORS	Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.		
TITLE	Medicago truncatula BAC Clone mth2-9d1		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 107857)		
AUTHORS	Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-AUG-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
COMMENT	----- Genome Center Center: Department Of Chemistry And Biochemistry The University Of Oklahoma Center code:UOKNOR ----- * NOTE: This is a 'working draft' sequence. It currently * consists of 9 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * 1 3029: contig of 3029 bp in length * 3030 3129: gap of unknown length * 3130 7573: contig of 4444 bp in length * 7574 7673: gap of unknown length		

```

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (27-FEB-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
5 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (03-MAR-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
6 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (16-APR-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
7 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (16-APR-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
8 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (26-AUG-1997) Plant Gene Expression Center, 800 Buchanan
St., Albany, CA 94710, USA
9 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (16-SEP-1997) Plant Gene Expression Center, 800 Buchanan
St., Albany, CA 94710, USA
On Aug 26, 1997 this sequence version replaced gi:1541948.
The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1.
FEATURES
Source
1. .110514
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F7G19"
971..2850
/genes="F7G19.1"
Join(971..1279,1540..1641,1688..1837,2134..2230,
2556..2665,2740..2850)
/genes="F7G19.1"
/notes="Similar to ATP-dependent Clp protease (gb|D90915).
EST gb|N65461 comes from this gene."
/codon_start=1
/evidence-not_experimental
/protein_id="AA070396.1"
/db_xref="GI:2342674"
/translation="MASCLQASNNLSLPRSSPSPPHPLSSSSGRRLKTPRYAFRA
KASAKTPMPPIPNKOPFLSTLASIAANSPEKLNRPVNADVPPYLDIPDSQMLWSPA
QVRSVAYNEHPRTPPPLPSYLLDGRIVYIGMPYDDIFCFQVPAVELVVAELM
YIQMLDKPPIYIINSTGTRDDGETVTCVGAAGACGLLSAGTKGRKMPHHA
KEESFCYFIFLIPWQLQVITNRDILVELLSKHTSGNSVETVANYMRRPYNDAPKA
KFGVTRVSLGVIP"
complement(3700..3771)
/products="tRNA-Glu"
/notes="codon recognized: GAA"
3942..6513
/genes="F7G19.2"
Join(3942..4046,4391..4507,4747..4854,4973..5164,
5263..5454,5551..5669,5747..5834,5916..6035,6216..6356,
6436..6513)
/genes="F7G19.2"
/notes="Similar to probable Mg-dependent ATPase
(pir|S56671). ESTs gb|T46782.gb|AA04738 come from this
gene."
/codon_start=1
/evidence-not_experimental
/protein_id="AA070397.1"
/db_xref="GI:2342675"

/translation="MAEDTSFEGDLASMTTDDIGRASRLIANRILKERSQRTNLD
LESVKEIKENQKIKLNPOLPYLVGNVILEMSPEDDAEDGANDLDSQKRGKV
VLTSRTQITFLPVGLAVSDTLKPGLVGNKDSYILDITLPSYDSRVKAMVDEK
PTDYNDIGGLEKIQELVEAIIVLPMTHKQPEKLGIRPKGHVLLGPPGTGKIMAR
ACAAQNTATKLKLAGPOLVQXFIGDGAKLVDAFLAKRSPCLIFIDEITAGTKRF
DSVSGDREVORTMLELNLQDGFSDDRIKVIAATNRADILDALMRSGRLDKRIEF
PHPTFARGRIQIHRSKMYNADVNFELARSTDDFNGAQKAVCEAGMALALRERDA
TEVNHEDFNEGIIQVQAKKASLNYYA"
complement(6693..9149)
/genes="F7G19.3"
Join(6693..6802,6892..7391,7471..7544,
7619..7721,7805..7920,8032..8127,8253..8639,8799..9149)
/genes="F7G19.3"
/notes="Strong similarity to Oryza NADPH oxidase
(gb|X93301)."
/codon_start=1
/evidence-not_experimental
/protein_id="AA070398.1"
/db_xref="GI:2342676"
/translation="MLSQKLPDKRNPVKRFAMNISYFFLENNKRIWLTWISICI
TLFTKFLQYKRTVFVEMVCYCVTKNGSAETLKNMALILLPVCRNTITWLTQSKL
IGSVPEDDNINPHKVVAGCIAGVIGLHAISHLACDFRLLHAKNVEFEPMKKEFGE
RPNYGFWMKGTGTCVTVMVLAIVLAQWFRNRANLPKSLRGLTGAFNFWYS
HELFEVIVYLLIVHGYVFLSKWEYHKTMTYLAVPVLYAFELIEAFRPGAKAVY
LKVAVPGNVLISLYMSKPKFKYTSQGYIYINCSDVPLQWHPFISASGDDYLVH
IRLGDWTSQKSLYSKVCOLPSTQSGLFIADIGQANNITRPRLLIDGPGYAPAO
YRNYDVLLVGLGIGATPLSIIRDVLNNIKNONSIERGTNOHIKNVATKRAYFYVY
TRQSGLEWPFSEVMNVAEYDSGMIEHLNYCTSVYEGDARSALITMLKSLHAKSG
IDVSGTRVTFHFRNWRVSFAHVAHVNVQVGVFCGTCGTCIIGLKLRLAQFSRK
TTTKFHFHKEF"
complement(9310..10430)
/genes="F7G19.4"
Join(9310..9492,9582..9630,9745..9904,
10061..10430)
/genes="F7G19.4"
/codon_start=1
/evidence-not_experimental
/product="F7G19.4"
/protein_id="AA070399.1"
/db_xref="GI:2342677"
/translation="MESSSGETNKISRCATGSDNDPEDEYVEITLVRDETINMKZA
KATLRSVLGRLKTMVKUSFSASRLDQSKSPGAFALGRFLAKNDAGRGWDEVA
KATPDLKLVGKPKPKGHCIGVSESEFYNELFEALVRRTGTTSSITKTELFPE
QITGNSFDRLQIIFDMVVDKLDGRITGDEVKEIILASASANKLSKIKENWDEVAALI
MEELDNLGYIEVRCLLSNGYTCULTSLKTLNP"
11553..13929
/genes="F7G19.5"
Join(11553..11812,11908..12119,12289..13011,13383..13344,
13419..13929)
/genes="F7G19.5"
/notes="Similar to Arabidopsis luminal binding protein
(gb|D89342)."
/codon_start=1
/evidence-not_experimental
/protein_id="AA070400.1"
/db_xref="GI:2342678"
/translation="MSSLAIEGEQKLTGVIDLGTYSQGVYHKKHVIANDQG
NRITPSWAPTDTERLIGRAAKNOAKNPRTIFDPKRLIGRFPDPPDVORDIKFLPY
KVUNKDKPYIOVKVKEGKLSPEISAMILTMEKTEAEFLGKKIKDAVITVTPAE
NDAQOATKADAGALAGNVVRIINEPTGAIAGLCKGGSNLLYDLGGGTFDVI
LTIDNGVFEVLSGTGTHLGGDFHVRMDYFIKLVKKYKNDIKSDHKAIGKLRREC
ELAKLSNQHQVRVEIESLFDGDFSEPLTRARFEELNMDLFSKTMKALKADAG
LKKSIDETVLVGGSTRIPKVQQLKDFDGPSPKGTNPDEAVYGAAYQVGVLSGE
QGEETOKNLTLLKTDILLDVAPLSLGIIETVGGVMTNIPRNTVITPKTSQVFTYD
QQTVTINVEGERSMTKDNRELKGFOLGILPAPRQVYRGVQLEVFTEVDANGIL
QVAKDKVAKTSQITITNDKRLTEEEHEMIREAREABEDKIKWEKIDARKKLET
YVYNNKSTVADKELKAKISDEDEKQGVGKLEALEWLENVNAEKEDYDEKLEKVEL
VCDPVYKVEITGEENEDTMTMTSYRIV"
complement(14446..15420)
/genes="F7G19.6"
/genes="F7G19.6"
complement(14446..15420)
/genes="F7G19.6"
/notes="Similar to Glycine SRC2 (gb|AB000130). ESTs

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 01:16:27 ; Search time 714 Seconds
(without alignments)
9376.975 Million cell updates/sec

Title: US-09-857-525C-1
Perfect score: 1576
Sequence: 1 gcacgagggaacatttcgcg.....tcaaaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1576	100.0	1576	3	AAA51347
2	987.4	62.7	2149	3	AAA51348
3	979.4	62.1	1632	3	AAA51353
4	594.2	37.7	2538	3	AAA51351
5	547.2	34.7	2160	3	AAA51350
6	529	33.6	538	3	AAA51354
7	266.2	16.9	282	6	ABL70761
8	259	16.4	514	3	AAA51357
9	148.4	9.4	702	3	AAA51349
10	139	8.8	178	6	ABL75404
11	95	6.0	542	3	AAA51355
12	50.6	3.2	1905	6	ABE551409
13	50.6	3.2	3261	4	ABL18857
14	50.6	3.2	39651	4	ABL18856
15	49.6	3.1	969	7	ACG60984
16	46.8	3.0	472	8	ACH36558
17	43.8	2.8	498	3	AAA51356
18	43.8	2.8	1266	3	AAA51352
19	43	2.7	2000	7	ADA71938
20	42.5	2.7	324	7	ABT22888
21	42.5	2.7	1576	2	AA78927
22	42.4	2.7	1578	6	ABL48616
23	42.4	2.7	1578	6	ABK84324

24	42.4	2.7	1578	7	ACA56515	Aca56515 Human sig
25	42.4	2.7	1691	3	ACA98774	Aac98774 Human pan
26	42.4	2.7	1774	6	ABL39755	Abi39755 Human NS
27	42	2.7	735	4	AAH07561	Aah07561 Human CDN
28	42	2.7	1167	6	AAH48614	Aal48614 Human ins
29	42	2.7	1435	2	AAH06882	Aax06882 Disease a
30	42	2.7	2000	7	ADA71938	Ada71938 Rice gene
31	42	2.7	2223	4	AAH14839	Aal14839 Human CDN
32	42	2.7	2223	6	AAH48615	Aal48615 Human ins
33	41.4	2.6	447	9	ADB57780	Adb57780 Toxicity-
34	41.4	2.6	990	6	ABQ69317	Abq69317 Listeria
35	41.4	2.6	996	6	ABQ67749	Abq67749 Listeria
36	41.4	2.6	81905	6	ABQ69244	ABQ69244 Listeria
37	41.4	2.6	82689	6	ABQ67198	ABQ67198 Listeria
38	41	2.6	491	9	ADE81890	Ades1890 Arabidops
39	41	2.6	2804	6	ABK35928	Abk35928 cDNA sequ
40	40.2	2.6	670	4	AAH26003	Aah26003 Human CDN
41	40.2	2.6	670	7	ABX73344	Abx73344 Human nov
42	40	2.5	375	7	ABX49849	ABX49849 Bovine ES
43	39.6	2.5	843	5	ABV16074	Abv16074 Human pro
44	39.4	2.5	459	6	ABL93683	Abi93683 Arabidops
45	39.4	2.5	594	3	AAC79823	Aac79823 Human sec

ALIGNMENTS

RESULT 1

AAA51347
ID AAA51347 standard; cDNA; 1576 BP.

XX
AC AAA51347;

DT 26-SEP-2000 (first entry)

DE Z. mays sucrose non-fermenting 4 protein cDNA from clone csk1c.pk001.c15.

XX Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
XX carbon catabolite repression; development; nitrogen partitioning; ss.

OS Zea mays.

XX Key Location/Qualifiers

FT CDS 3..1358

FT /*tag= a

FT /label= SNF4

FT /product= "Sucrose non-fermenting_4_protein"

FT /partial

XX WO200036116-A2.

PD 22-JUN-2000.

XX 15-DEC-1999; 99WO-US029825.

XX 16-DEC-1998; 98US-0112564P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Allen SM, Heppard EP, Miao G, Weng Z;

XX WPI; 2000-431594/37.

XX P-PSDB; AAY96783.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
PT involved in carbon catabolite repression in plants and seeds; useful for
PT controlling carbon and nitrogen partitioning pathways during plant growth
PT and development.

XX Claim 2; Page 31; 48pp; English.

XX This cDNA encodes maize (Zea mays) sucrose non-fermenting 4 (SNF4)

CC protein which is involved in carbon catabolite repression in plants and

CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
 CC cskic.pk001.c15 prepared from corn unpollinated developing silk 24 hours
 CC after emergence. The polynucleotides are used in plants to control carbon
 CC and nitrogen partitioning pathways during plant growth and development.
 CC The catabolite repression proteins would facilitate studies for better
 CC understanding the mechanism of catabolite repression in plants and could
 CC provide genetic tools to enhance or otherwise alter the accumulation of
 CC carbohydrates, lipids and proteins during plant growth and development
 XX
 SQ Sequence 1576 BP; 418 A; 336 C; 362 G; 460 T; 0 U; 0 Other;

Query Match 100.0%; Score 1576; DB 3; Length 1576;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGGGAACATTTGCCGATGTCCTCTATCGAAGGCTGCCCACTGTAATTCAGGCTA 60
 DB 1 GCACGAGGGAACATTTGCCGATGTCCTCTATCGAAGGCTGCCCACTGTAATTCAGGCTA 60

QY 61 TTGCGAGTCGTCTCTCGGAGTTACAGAGTACAAATTTCTTGTGAGCGGGAAATGGCGC 120
 DB 61 TTGCGAGTCGTCTCTCGGAGTTACAGAGTACAAATTTCTTGTGAGCGGGAAATGGCGC 120

QY 121 ATGATGAGCGTCAACCTTACCATATCTGGGAGTTTGGCATAGTTAAACACACTTTACTTGA 180
 DB 121 ATGATGAGCGTCAACCTTACCATATCTGGGAGTTTGGCATAGTTAAACACACTTTACTTGA 180

QY 181 CAAGGGAATTAACCAATAAACAACCTTATCAAGTCCAAAGCACACCTCGAAGCAGGATGA 240
 DB 181 CAAGGGAATTAACCAATAAACAACCTTATCAAGTCCAAAGCACACCTCGAAGCAGGATGA 240

QY 241 ACATGAGTGGTAATTAAGAAATTTTCAAGTACGTTAGTCTGAGTGGCACCGT 300
 DB 241 ACATGAGTGGTAATTAAGAAATTTTCAAGTACGTTAGTCTGAGTGGCACCGT 300

QY 301 CAGAGGACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTCTG 360
 DB 301 CAGAGGACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTCTG 360

QY 361 AATATCTGAATTTGCATACATGCTATGATTTACTCCAGATCTCGGCAAGGTTATTCGCC 420
 DB 361 AATATCTGAATTTGCATACATGCTATGATTTACTCCAGATCTCGGCAAGGTTATTCGCC 420

QY 421 TAGACATTAATTTACCTGTGAAGCAATTCATTCATATCTCATGATGACAGGGGATTCCTG 480
 DB 421 TAGACATTAATTTACCTGTGAAGCAATTCATTCATATCTCATGATGACAGGGGATTCCTG 480

QY 481 TAGCTCTCTCTGGGACTCATTCAGAGTCAATTTGTTGCTCTCTTAGCCCATTTGGATT 540
 DB 481 TAGCTCTCTCTGGGACTCATTCAGAGTCAATTTGTTGCTCTCTTAGCCCATTTGGATT 540

QY 541 TCATCTCATATTTGGGGAGCTAGAACTCATGCTCGAATTCACAGAGCAGCTTG 600
 DB 541 TCATCTCATATTTGGGGAGCTAGAACTCATGCTCGAATTCACAGAGCAGCTTG 600

QY 601 AAACACACATATATCTGATGGAAGAGGCTAAGCGCAAACTAATCGAAGAATGATA 660
 DB 601 AAACACACATATATCTGATGGAAGAGGCTAAGCGCAAACTAATCGAAGAATGATA 660

QY 661 GTGAGTGGCGGACCGCAACAGCATCTAGTGCATGTCACCCCTTATGAGTCTTGAAGGACA 720
 DB 661 GTGAGTGGCGGACCGCAACAGCATCTAGTGCATGTCACCCCTTATGAGTCTTGAAGGACA 720

QY 721 TTGAGTAAAGCTTTTGAAGTGGCAATTTACAGTGGCAGTATTTATTCATCATCAT 780
 DB 721 TTGAGTAAAGCTTTTGAAGTGGCAATTTACAGTGGCAGTATTTATTCATCATCAT 780

QY 781 CAGATGGAATCATTCGCGAGTTATTCGATCTTGCATCTTCTGGAATTTGAAATGTA 840
 DB 781 CAGATGGAATCATTCGCGAGTTATTCGATCTTGCATCTTCTGGAATTTGAAATGTA 840

QY 841 TTGATGATATCTTCAAAACTCAACTGGTAATTTGCTTATCTGAAACCAACGAGTGTGCT 900

DB 841 TTGATGATATCTTCAAAACTCAACTGGTAATTTGCTTATCTGAAACCAACGAGTGTGCT 900

QY 901 CCATTCGCTGGGTTCTGGGTTCCGAAATTTGGTATCTGAAACAGTCTGTCATTTGGCTA 960

DB 901 CCATTCGCTGGGTTCTGGGTTCCGAAATTTGGTATCTGAAACAGTCTGTCATTTGGCTA 960

QY 961 TGTGCGACCTTAATGCTCACTTAGCTCTGCCCTTAACATGTTGTTCAAGCTCGAGTAA 1020

DB 961 TGTGCGACCTTAATGCTCACTTAGCTCTGCCCTTAACATGTTGTTCAAGCTCGAGTAA 1020

QY 1021 GCTCAATACCAATTTGTGGATGACAAACGACTCCCTGCTTGACACTTACTCTAGAAGTGACA 1080

DB 1021 GCTCAATACCAATTTGTGGATGACAAACGACTCCCTGCTTGACACTTACTCTAGAAGTGACA 1080

QY 1081 TCACAGCCCTAGCAAAAGACAAAGTCTACACATGTTGCTGGTGGATGAGATGACCAATTC 1140

DB 1081 TCACAGCCCTAGCAAAAGACAAAGTCTACACATGTTGCTGGTGGATGAGATGACCAATTC 1140

QY 1141 ACCAGGCTTTGACGCTTGGACAAAGATGCAATACACCTTTTGGATTTCTTAAACGCCAGA 1200

DB 1141 ACCAGGCTTTGACGCTTGGACAAAGATGCAATACACCTTTTGGATTTCTTAAACGCCAGA 1200

QY 1201 GATGCGAGATGTCCTCCGCTCTGATCTTTGCTGAAGTGTATGGAGCGACTGGCTAATC 1260

DB 1201 GATGCGAGATGTCCTCCGCTCTGATCTTTGCTGAAGTGTATGGAGCGACTGGCTAATC 1260

QY 1261 CTGGGGTGGCGCGGGTGTCTCAATGTTGGAAGCTGGAGCAACCGTGGAGGCGCATCATAT 1320

DB 1261 CTGGGGTGGCGCGGGTGTCTCAATGTTGGAAGCTGGAGCAACCGTGGAGGCGCATCATAT 1320

QY 1321 CACTAAGTGACATTTTCAAGTTCCTTGTGAGCTTTGTGAGGAGAGGAGCATGCGGAGT 1380

DB 1321 CACTAAGTGACATTTTCAAGTTCCTTGTGAGCTTTGTGAGGAGAGGAGCATGCGGAGT 1380

QY 1381 TTTTAAACAGTAGCAACCTCTGGGCTCTTGGCTTGGTCTGCGGATATTTGCTATCTGTC 1440

DB 1381 TTTTAAACAGTAGCAACCTCTGGGCTCTTGGCTTGGTCTGCGGATATTTGCTATCTGTC 1440

QY 1441 AGGCACTCNGGTTTCATGCTCTGTAAGTCTCTTTTGTGTTTGTGTTTACCATTT 1500

DB 1441 AGGCACTCAGGGTTTCATGCTCTGTAAGTCTCTTTTGTGTTTGTGTTTACCATTT 1500

QY 1501 GGCCTCCCTCGGTTTAAACAAACATACATCAGGTGATTTTATCTATTGATTTGTTCAA 1560

DB 1501 GGCCTCCCTCGGTTTAAACAAACATACATCAGGTGATTTTATCTATTGATTTGTTCAA 1560

QY 1561 AAAAAAAAAAAAAA 1576

DB 1561 AAAAAAAAAAAAAA 1576

RESULT 2

AAAS1348

ID AAAS1348 standard; cDNA; 2149 BP.

XX

AC AAAS1348;

XX

DT 26-SEP-2000 (first entry)

XX

DE Rice sucrose non-fermenting 4 protein cDNA from clone r10n.pk083.i18.

XX Rice; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

KW carbon catabolite repression; development; nitrogen partitioning; ss.

XX

OS Oryza sativa.

XX

Key Location/Qualifiers

FT CDS 185..1666

FT /*tag= a

FT /product= "Sucrose_non-fermenting_4 protein"

XX

PN WO200036116-A2.

```

XX PD 22-JUN-2000.
XX PF 15-DEC-1999; 99WO-US029825.
XX PR 16-DEC-1998; 98US-0112564P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E. I.
XX PI Allen SM, Heppard EP, Xiao G, Weng Z;
XX DR WPI; 2000-431594/37.
XX DR P-PSDB; AAY96784.
XX PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
XX PT involved in carbon catabolite repression in plants and seeds, useful for
XX PT controlling carbon and nitrogen partitioning pathways during plant growth
XX PT and development.
XX PS Claim 2; Page 33-34; 48pp; English.
XX CC This cDNA encodes rice (Oryza sativa) sucrose non-fermenting 4 (SNF4)
XX CC protein which is involved in carbon catabolite repression in plants and
XX CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
XX CC Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
XX CC r10n.pk083.118 prepared from rice 15 day old leaf. The polynucleotides
XX CC are used in plants to control carbon and nitrogen partitioning pathways
XX CC during plant growth and development. The catabolite repression proteins
XX CC would facilitate studies for better understanding the mechanism of
XX CC catabolite repression in plants and could provide genetic tools to
XX CC enhance or otherwise alter the accumulation of carbohydrates, lipids and
XX CC proteins during plant growth and development
XX SQ Sequence 2:49 BP; 523 A; 492 C; 501 G; 633 T; 0 U; 0 Other;

Query Match 62.7%; Score 987.4; DB 3; Length 2149;
Best Local Similarity 83.2%; Pred. No. 6.9e-266;
Matches 1123; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 5 GAGGGAACATTTCGGATGTCCTATCGAAGCTGCCACCTGTATTTTCAGGCTATTG 64
DB 316 GACAGAACAATTTCGCAATGTCCTGTTGAAGTTGTCCGCCGTTTTTCAGGCCATATG 375
QY 65 CAGTCTGTCTCTCGGATTCACAGATCAAAATCTTTTGGACGGGAATGGCGCATGA 124
DB 376 CAGTTTGTCCCTGGGATTCATCAGTACAAATTTTGTGGATGGGAATGGCGCACGA 435
QY 125 TGAGCGTCAACCTACCATATCTCGGAGTTTGGATAGTTTACACACTTTACTTGTCAAG 184
DB 436 TGAGCGGCAACCTACCATATAACAGGAGACTATGGTGTGTTTATGCTTGACTAG 495
QY 185 GGAATATAACCAATAAACAACCTTATCAAGTCCAAAGCACACCTTGGAGAGGATGAACAT 244
DB 496 GGACTTTGACGATTAATTAACAATTTAGCCCTAGTACACCTTGGAGTAGATGAACAT 555
QY 245 GGATGTGATATGAAATTTTCAACGTTACGGTTACGTTTACAGTGGGACCGTTTCAGA 304
DB 556 GGATGTGGACACGACAAATTTTCAACGTTACTGTTCTTCTGCTGATGGCAATTTATCAGGA 615
QY 305 AGGTACTCTGAGATTTTCAGAGGCTGCATACAAATATCTAGGTGCTGTTCTGATATA 364
DB 616 AGGTCTCTCAGAAATTTTCAGAGGCGCTATACAGATCTCTAGGTGCTGTTAGCGGATTT 675
QY 365 TCTGAATTTGCATACATGCTATGATTTACTCCAGATTTCTGGCAAGTTATTCCTCCATGA 424
DB 676 TCTGAATGGCAAACTGGGTATGATTTACTCCAGATTTCTGGCAAGTATTCCTCTCAGA 735
QY 425 CATTATTTTACCTGTGAAGCAATCAITTCATATTTCTCCATGAAACAGGGGATTCCTGTAGC 484
DB 736 CGTTAATTTCCCTGTGAAGCAATCTTTTCATATTTCTCATGAAACAGGGAATTCCTGTGGC 795
QY 485 TCTCTCTGGACTCAITTCAGAGGTCATTTGTTGCTCTCTCTAGGCCATTTGATTCAT 544

```

RESULT 3
 AA51353
 ID AAA51353 standard; cDNA; 1632 BP.
 XX
 AC AAA51353;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE Wheat sucrose non-fermenting 4 protein cDNA from clone wll.pk0002.b3.
 XX
 KW Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX carbon catabolite repression; development; nitrogen partitioning; ss.
 OS *Triticum aestivum*.

XX FH Key Location/Qualifiers
 XX FT 10.1393
 XX FT /*tag= a
 XX FT /product= "Sucrose_non-fermenting_4_protein"
 XX PN WO200036116-A2.
 XX PD 22-JUN-2000.
 XX PF 15-DEC-1999; 99WO-US029825.
 XX PR 16-DEC-1998; 98JS-0112564P.
 XX PA {DUPO } EU PONT DE NEMOURS & CO S I.
 XX PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX DR WPI; 2000-431594/37.
 XX PS P-PSDB; AAY96789.
 XX PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX PS Claim 18; Page 43; 48pp; English.
 XX PS This cDNA encodes wheat sucrose non-fermenting 4 (SNF4) protein which is
 CC involved in carbon catabolite repression in plants and seeds. The cDNA
 CC was isolated, based on similarity to SNF4 proteins from Arabidopsis
 CC thaliana and Saccharomyces cerevisiae, from library will prepared from
 CC wheat leaf from 7 day old seedling. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX SQ Sequence 1632 BP; 431 A; 370 C; 389 G; 442 T; 0 U; 0 Other;

Query Match 62.1%; Score 979.4; DB 3; Length 1632;
 Best Local Similarity 83.5%; Pred. No. 1.1e-263;
 Matches 1125; Conservative 3; Mismatches 216; Indels 6; Gaps 1;
 QY 12 CATTTGGCGATGCTCTCTATGAGGCTGCCCCACTGATTTTCAGGCTATTTCAGTCTG 71
 DB 2 CACGAGCCGATGCTCCAGTGGAGGTGCCCCACTGATTTTCAGGCTATTTCAGTCTG 61
 QY 72 TCTCCTGGGATTCAGAGTACAAATTTCTTTGGACGGGAAATGGCGCATGATGAGCGT 131
 DB 62 CTTCCAGGATTTATCAGTACAAATTTCTTTGGACGGGAAATGGCGCATGATGAGCGG 121
 QY 132 CAACCTACCATATCTGGGGAGTTGGCATAGTTAAACACTTTTACCAAGGGAATAT 191
 DB 122 CAACCTACTACTATCTGGAGAGTATGGGGTGGTAAACACTTTTATATCTTCAACAGGGAAT 181
 QY 192 AACCAATAAACACTTTATCAAGTCCCAAGCACCTCGAACGAGATGACATCGATGTG 251
 DB 182 GACCACATAAATCTACTGACGCCCCACCTACCTGGGAGCAGG-----ATGGATGTG 235
 QY 252 GATAATGAAATTTTCAACCTACGTTACGTTGTCAGATGSCACCGTTTCAGAGGTACT 311
 DB 236 GACAGTACAGTTTTCACGAAATGGGTTGTTGTCGATGTTGCCCTTCAGGAGGTCT 295
 QY 312 CTGAGAGTTTCAGAGGCTCAATACAAATATCTAGTGTGTTTCTGAAATATCTGAAT 371
 DB 296 CCAAGAAATCTCAGAGGCTGCTATACAGATCTCTAGTGTGTTGTTGCTGAGTATCTGAAT 355
 QY 372 TCGATACATGCTATGATTTTACTCCAGATTTCTGGCAGGTTATTCCTAGACATTAAT 431
 DB 356 GCGCATACAGGCTATGACCTACTACAGATCTTGGAAAGGTCAITGCTCTGGACATTAAT 415

QY 432 TTACCTGTGAAGCAATCAITTCATATTTCTCCATGAACAGAGGATTCCTGTAGCTCTCTC 491
 DB 416 TTACCTGTGAAGCAATCTTTTCCATATTTCCATGAACAGAGGATTCCTGTAGCTCTCTC 475
 QY 492 TGGGACTCATTGAGAGTCAATTTGTTGTTCTCTCTAGCCCATTTGATTTTCACTCATTA 551
 DB 476 TGGGATTCATTGAGGGGTGAGTTTGTGGCTTTCTGAGCCCACTGGATTTTATCTTATA 535
 QY 552 TTGGGGAGCTAGAACTCATGCTCGAACTTACACAGAGAGAGGAGCTTGAACACACACT 611
 DB 536 TTGAGAGAGCTGAACTCATGCTCAAACTTACAGAGAGAGAGCTTGAACACACACT 595
 QY 612 ATATCTGATGAGAGAGCTTAAGCGGCAAACTTAATGGAAGAAATGATAGTCACTAGTGGCG 671
 DB 596 ATATCTGCTGGAAGAGAGCTTAAGCGGCAAACTTAATGGAAGAAATGATAGTCACTAGTGGCG 655
 QY 672 CCGCAACAGATCTAGTGCATGCCACCCCTTATGAGTCTTGTAGGAGCAATGTCAGTAAAG 731
 DB 656 TCAATCAGCATCTAGTGCATGCCACCCCTTATGAGTCTTGTAGGAGCAATGTCAGTAAAG 715
 QY 732 CTTTTCGAAATGCGCATTTCTACAGTCCAGTTTATTTTTCATCATCATCATCATCATCA 791
 DB 716 ATACTCGAACTGCGCATTTCTACAGTCCCAATCTCTTTCATCATCATCATCATCATCAT 775
 QY 792 TTCCCGCAGTTATGCACTTGCATCTTCTGCACTTTCTGGAATTTTGAATATGTTTGTAGTAC 851
 DB 776 TTTCGCGAGCTTGCATCTTGCATCTTCCAGGAAATTTTGAATATGTTTGTAGTAC 835
 QY 852 TTCAAAATCAACTGGAATTTGCTATTTGAAACCAACAGTGTGCTTCCATTCGCTG 911
 DB 836 TTCAAAATCAACTGGAATTTGCTATTTGAAACCAACAGTGTGCTTCCATTCGCTG 895
 QY 912 GGTCTCTGGTTCGAAATTTGCTATCTGAAACAGTGTGCTTGTGCTATTTGTGACCT 971
 DB 896 GGTACTGGTTCGAAATTTGCTATCTGAAACAGTGTGCTTGTGCTATTTGTGACCT 955
 QY 972 AATGCTCATTAGTCTGCTGCTTAAACAGTGTGCTTGTGCTATTTGTGACCT 1031
 DB 956 AATACATCTTGTAGCTTGTGCTTAACTTGTGCTTGTGCTATTTGTGACCT 1015
 QY 1032 ATTGTGATGACAGCTCTGCTGCTTGTGCTATTTGTGCTATTTGTGACCT 1091
 DB 1016 ATTGTGATGATGACAGCTCTGCTGCTTGTGCTATTTGTGCTATTTGTGACCT 1075
 QY 1092 GCAAAAGCAAGGTCTACACCATTTTGGCTGATGAGATGACCATTTCAACAGGCTTTG 1151
 DB 1076 GCGAAGCAAGGTCTACACCATTTTGGCTGATGAGATGACCATTTCAACAGGCTTTG 1135
 QY 1152 CAGCTTGGCAAGATGCAATACACCTTTTGGATTTTAAAGCGGAGATGCGAGATG 1212
 DB 1136 CAGCTCGGGCAAGATGCAATACACCTTTTGGATTTTAAAGCGGAGATGCGAGATG 1195
 QY 1212 TGGCTCGGCTGATCTCTTGTGCTGAGGATGAGGAGCTGCTTAACTCTGGGCTGGG 1271
 DB 1196 TGTCTCAGTCTGACCTTTGCTGAGGATTTGAGGATTTGGCTATTTCTTGGGCTGGT 1255
 QY 1272 CGGTGTTTCTTGTGAGGCTGGGAGCAAACTGTGGAGGCGCATCATATCACTAAGTGAC 1331
 DB 1256 CGGTGTTTCTTGTGAGGCTGGGAGCAAACTGTGGAGGCGCATCATATCACTAAGTGAC 1315
 QY 1332 ATTTTCAAGTTCTTGTGAGGCTTTGTA 1358
 DB 1316 ATATTCAGTTGCTGCTGAGCTAGCGA 1342

RESULT 4
 AAA51351
 ID AAA51351 standard; cDNA; 2538 BP.
 XX AAA51351;
 XX
 DT 26-SEP-2000 (first entry)

XX Soybean sucrose non-fermenting 4 protein cDNA from clone sfl1.pk0004.b4.
 DE
 XX
 XX
 KW Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.
 OS
 XX Glycine max.
 XX
 FH Key Location/Qualifiers
 FT CDS 395..1873
 FT /*tag= a
 FT /product= "sucrose non-fermenting_4_protein"
 XX
 XX WO200036116-A2.
 PD 22-JUN-2000.
 XX
 PF 15-DEC-1999; 99MO-US029825.
 XX
 PR 16-DEC-1998; 98US-0112564P.
 XX
 PA (DUF0) DU PONT DE NEMOURS & CO E I.
 XX
 PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX
 XX WPI: 2000-431594/37.
 DR P-PSDB; AAY96787.
 XX
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX
 XX Claim 2; Page 39-40; 48pp; English.
 PS
 XX This cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)
 CC protein which is involved in carbon catabolite repression in plants and
 CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from library sfl1
 CC prepared from soybean immature flower. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX
 SQ Sequence 2538 BP; 670 A; 463 C; 571 G; 834 T; 0 U; 0 Other;
 Query Match 37.7%; Score 594.2; DB 3; Length 2538;
 Best Local Similarity 66.8%; Pred. No. 1.2e-155;
 Matches 896; Conservative 0; Mismatches 433; Indels 12; Gaps 3;
 QY 8 GGAACATTGGCGATGTCCTCTATCGAAGCTGCCCCACCTGTAATTCAGGCTATTGGCAG 67
 DB 529 GGAGCTCTACCGATGTCGCCAGTGGAGAGTGTCCAACTGTGTTCAAGTGAATTAATA 588
 QY 66 TCTGTCTCTCGGATTCACAGTACAAATTCCTTGTGGAGCGGGAATGGCGCATGTA 127
 DB 589 CTGGACCCGGTTACCATCAGTACAAAGTTTGTGTGGAGAAATGGCGGATGATGA 648
 QY 128 GGCTCAACCTACCATATCTGGGAGTTTGGCATAGTAAACACACTTACTTGAACAGGGA 187
 DB 649 ACATCAACCTTATGTACCTGGAGAAATATGGATAGTCAACACTGTCTTATGGCCACTGA 708
 QY 188 ATATACCAAAATAAACCTTATCAAGTCCAGCACACCTGGAGCAGGAATGAATGGA 247
 DB 709 TCCTAACATACCTGCTTTCACCTCCAGACGTTGCTTCTGGAAATAG-----CATGGA 762
 QY 248 TGTGATAATGAAATTTTCAAGTACGGTTAGTTGTGAGTGGCACCGTTTCAGAAAG 307
 DB 763 TGTGGATAATGATGCTTTTGGCGGAATGGCCGGTTGACCGATGACTTTGAGTGAGGT 822
 QY 308 TACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGCTGTGTTTCTGAATATCT 367
 DB 823 GCTGCCAAGAAATACAGATACTGATGTACAAATATCCGTCAGCGTATTTCTGCAATTTCT 882
 QY 368 GAATTGCAATATGCTATGATTTACTCCAGATTCCTGGCAAGGTTATTCGCCCTAGACAT 427
 DB 883 ATCTTCAACACACCGCTTATGAATTACTCCGAGTCAGGCAAGGTTGTTCTGTTGGATGT 942
 QY 428 TAATTTACCTGTGAAGCAATCATTTCCATATCTCCATGACACAGGAGATTCCTGTAGCTCC 487
 DB 943 TGATCTACCAAGTGAACACAGCATTTTCATATATTCGATGAGCAGGAGTTTTCATAGGCTCC 1002
 QY 488 TCTCTGGGACTCATTCAGAGGTCATTTGTGTGCTCTCTTAGCCCATTCGATTTCACTACT 547
 DB 1003 TCTTTGGGACTTCTGCAAGGGCAATTTTGTGTGCTCTTAGTCTCTGATTTTATTTT 1062
 QY 548 CATATTGGGGAGCTAGAAACTCATGGCTGCAACTTGCACAGAGAGAGAGCTTGAACACA 607
 DB 1063 AATTTTAAGAGAGCTGGGGAATCATGATCCAACTCTGACAGAGAGAGAGCTTGAACACA 1122
 QY 608 CACTATATCTGCATGGAAGAGGCTTAAGCGGCAAACTTAATGGRAGAAATCATAGTCACTG 667
 DB 1123 TACCATATCAGCTTGAAGAGAGGAAATCGTATCTAAATAGACAGAACAAATGGACATGG 1182
 QY 668 --GCCACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTTGAGGAGCATTGC 724
 DB 1183 AACTGCAATTTCAAGATGTTTTATCCATGCAAGGCGCATATGTAATCTGAAAGATA-TGC 1242
 QY 725 AGTAAAGCTTTTGGAAAATGGCAATTTCTACAGTGGCCAGTTATTTATTCATCATCAGACA 784
 DB 1243 CATGAAGATCTTTGCAAAAGAGGTTTCAACTGTCTCTATTTATCCATTCATCTTCTGAAGA 1302
 QY 785 TGGATCAATCCCGCAGTATGTCATCTTCATCATCTTCCTGGAATTTTGAAGTATTTG 844
 DB 1303 TGTCTCATTTCCACAGTTACTACATCTGTCTTCACTTTCAGGAATCTTAAATGCAATTTG 1362
 QY 845 TAGATACTTTCAAAAACCTCAACTGGTAAATTTGCCCTATTTCTGAACCAACAGCTGTCTCCAT 904
 DB 1363 TAGTATTTTAGGCACTGCTCTAGTTCTTCTGCTGTACTTCAACTTCCAATCTGGCAAT 1422
 QY 905 TCGCTGGGTTCTCGGGTTCGAAATTTGGTGATCTGAAACAGTCTGTCATTTGGCTATGTT 964
 DB 1423 ACCGTGGGCACTGGGTGCCAAAATTTGGGGAATCAAAATCGCGGCGCTCTAGCAATGTT 1482
 QY 965 GCGACCTAATGCTCACTTAGTCTCTCCCTTAACATGTTGGTTCAAGCTGGAGTAAGCTC 1024
 DB 1483 GAGACCAACCGCTTCTCTTGGCTCAGCCCTAAATTTATTTAGTTTCAAGCCCAAGTAACTC 1542
 QY 1025 AATPACCAATTTGGATGACAAACGACTCCCTGTGTGACACTTACTCTAGAAGTGAACATCAC 1584
 DB 1543 AATACCAATAGTTGATGATAATGACTCTATTTTGGATATATATCTGCGGAGTGACATAAC 1662
 QY 1085 AGCCCTAGCAAAAGACAAGTCTACACAGATGTCGGCTGGATGAGTGAAGTACCATTACCA 1144
 DB 1603 AGCTTTGGCAAAAGAACAGAGCATATACATATTAATCTGTGCAAAATGACTGTTCATCA 1662
 QY 1145 GGCTTTGACGCTTGGACAAGATGCCAATACACCTTTTGGATTTCTTAAACGCCACAGATG 1204
 DB 1663 GGCTTTGAGTTGGGCGCAGGACGATATAGTCCCTATG---AGCTTAGAAGTCAAGATG 1719
 QY 1205 CCAGATGTGCTCCGGTCTGATCTTGTCTGAAGGTGATGGAGCGACTGGCTTAATCTCTGG 1264
 DB 1720 TCAGATGTGTTTGGCTTCTGATCTCTGCAATAAGTGAAGGAGCTTGGCAATTCAGG 1779
 QY 1265 GGTGCGCGGCTTCTCATTTGTGAAGCTTGGAGCAACCTGTGGAGGCGATCATATCACT 1324
 DB 1780 TGTGAGCGGCTTGTCTCATCTGGGAAGCTGGCAGCAAGCGGTGTAGAAGGCAATTTTCATT 1839
 QY 1325 AAGTGACATTTTCAAGTCTTT 1345
 DB 1840 GAGTGACATTTCAAGTCTTT 1860

RESULT 5
 ID AAA51350 standard; cDNA; 2160 BP.
 AC AAA51350;
 DT 26-SEP-2000 (first entry)
 XX Soybean sucrose non-fermenting 4 protein cDNA from clone ses4d.pk0040.b1.
 XX
 XX Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX carbon catabolite repression; development; nitrogen partitioning; ss.
 XX
 XX Glycine max.
 XX
 XX Key location/Qualifiers
 XX CDS 185..1633
 XX /*tag= a
 XX /product= "sucrose_non-fermenting_4_protein"
 XX
 XX WO200036116-A2.
 XX
 XX 22-JUN-2000.
 XX
 XX 15-DEC-1999; 99WO-US029825.
 XX
 XX 16-DEC-1998; 98US-0112564P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Allen SM, Heppard EP, Miao G, Weng Z;
 XX WPI; 2000-431594/37.
 XX P-ESDB; AAY96785.
 XX
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 XX involved in carbon catabolite repression in plants and seeds, useful for
 XX controlling carbon and nitrogen partitioning pathways during plant growth
 XX and development.
 XX
 XX Claim 2; Page 36-37; 48pp; English.
 XX
 XX This cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)
 XX protein which is involved in carbon catabolite repression in plants and
 XX seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 XX Arabidopsis thaliana and Saccharomyces cerevisiae, from library ses4d
 XX prepared from soybean embryogenic suspension 4 days after subculture. The
 XX polynucleotides are used in plants to control carbon and nitrogen
 XX partitioning pathways during plant growth and development. The catabolite
 XX repression proteins would facilitate studies for better understanding the
 XX mechanism of catabolite repression in plants and could provide genetic
 XX tools to enhance or otherwise alter the accumulation of carbohydrates,
 XX lipids and proteins during plant growth and development
 XX
 XX Sequence 2160 BP; 588 A; 392 C; 485 G; 693 T; 0 U; 2 Other;
 XX
 XX Query Match 34.7%; Score 547.2; DB 3; Length 2160;
 XX Best Local Similarity 66.0%; Pred. No. 1.6e-142;
 XX Matches 890; Conservative 0; Mismatches 413; Indels 45; Gaps 5;
 XX
 XX 3 GACATTTGGCGATGCTCTCTATGGAAGGCTGCGCCACTGTATTTAGGCTATTTCAGT 68
 XX 317 GAACTTCTACAAATGTGCGCTTGGGAAGGCTGCGCCACTGTGTTTCAAGTTATTTCATAGC 376
 XX 69 CTCTCTCTGGGATTCACGAGTACAATCTTTTGGGACGGGGAATGGCGCATGTAG 128
 XX 377 TTGGTACCTGGTCATCATCATGATACAGTTTCTTTGTTGATGGAGATGGCGCATGATGAC 436
 XX 129 CGTCAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACACTTTTACCTTTGACAAAGGAA 188
 XX 437 CTTCAACCTTGTGAATCTGGGAATATGGATGTTTAACTGTTTCACTGTTTCACTGTTGCTATCAT 496
 XX 189 TATAACCAATAAACACTTTATCAAGTCCAAAGCACACCTTGGAAAGCAGGATGAACATGGAT 248

Db 497 CCTAATATTTTACCTGTTTAACTCCAGACATAGTTTCTGGA-----CTAATACATGGAT 550
 QY 249 GTGGATAATGAAATTTTCAACGTACGGTTAGTGTGTCAGATGGACACCGTTTCAGAAAGT 308
 Db 551 GTGCAACACGAGGCTTTTCGACGCAATGGTTCGATTTGACAGATGATTAATGATG 610
 QY 309 ACTCTG---AGAGTTTCAGAGCTGCAATACAAATATCTAGTGTGCTGTTCTGAATAT 365
 Db 611 TATTTGCCAAGAATATCTGATTTGATATACAGACTCTCGTCAAGCTATTTCTGCAATTC 670
 QY 366 CTGAATTTGCTATACATGCTATGATTTACTCCAGATTTCTGCAAGTGTATTCGCCCTAGAC 425
 Db 671 CTATCTATGAGTACAGCATATGAAATTTCTCTGAGTCAGGCAAGTTGTTTACCTTGGAT 730
 QY 426 ATTAATTTACCTGTGAAGCAATCATTTCCATATTTCTCCATGAACAGGGATTTCTCTAGCT 485
 Db 731 GTTGATCTGCGCTGTGAACACAGCAATTTCCATATCTGATGATGACAGGAAATTTCCCATTC 790
 QY 486 CCTCTCTGGGACTCATTTGAGAGGTTCAATTTGTTGTTCTCTCTAGTCCCTTGGATTTCTATA 545
 Db 791 CCTCTCTGGGACATCTGCAAGGGGCGATTTGTTGGAGTTCTTAGTGCCTTGGATTTTATTT 850
 QY 546 CTCATATTTGCGGAGCTAGAAACTCATGCTCGAACTTTGACAGAAGACAGCCTTCAAAACA 605
 Db 851 TTAATTTTAAGAGAGCTCGGAATCATGGTCCATCTTAACAGAAGAGAGCTTGAACA 910
 QY 606 CACATATATCTGCATGGAAGAGAGCTTAAGCGGCAAACTTAATGGAGAAATGATAGTCAG 665
 Db 911 CATACCATATCAGCTTGAAGAGGAGAAATGAGACA----- 946
 QY 666 TGGGACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCTTTCAGGGACATTCGA 725
 Db 947 ---GGATTTACAAATGTTTATCCGTGCGAGGGCCATACGATAATTTGAAAGAGATTCCT 1003
 QY 726 GTAAAGCTTTTGCAAAATGGCAATTTCTACAGTGCCAGTTTATTTATTCATCATCATCAT 785
 Db 1004 GTGAAGATCTCTGCAACATGGAATTTCAACAGTCTCTATTTATTCATCAGA-----AGAT 1057
 QY 786 GGATCATTCGCGCAGTTTATGCAATCTTTCATCTGCAATCTTTCGAAATTTTGAATGTTTGT 845
 Db 1058 GGTTCATTTCCACAGTACTACATCTTTCATCTTTCAGGAATCTCTTAAATGCAATTTGC 1117
 QY 846 AGATACCTTCAAAACTCAACCTGTAATTTGCTATTTGACCAACCAAGCTGCTCCATTT 905
 Db 1118 AGTATTTTGAAGATGTTTCTAGTTCATGCTCTATCTTCACTTCACTTCACTTTCGCAATC 1177
 QY 906 CCGCTGGGTTTCTGGGTTCCGAAAATTTGGTATCTGAACAGTCTCTCATTTGGCTATGTTG 965
 Db 1178 CTTGTGGGACGTTGGTGCCCAAAATTTGGGAGTCAATCGCGGCTCTAGCAATGTTG 1237
 QY 966 CGACCTAATGCTCCTACCTAGTCTGCGCTTAACTGTTGGTTCAAGCTGGAGTGAAGTCA 1025
 Db 1238 AGACCAAAATGCTTCACTTCTTCTAGTCTTAACTTATTTAGTTCAGTCAAGTAAAGTTCA 1297
 QY 1026 ATACCAATTTGGATGACACAGCTCTCTGCTGACACTTACTCTAGAAGTGACATCACA 1085
 Db 1298 ATACCAATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1357
 QY 1086 GCCTTAGCAAAAGACAGGTTCTTACACATGTTTGGCTGGATGAGATGACCATTTCCACAG 1145
 Db 1358 GCTTTGGCAAGGACAGAACTTATCCCATATTAATCTTGTGATGAATGATGATGATGATGAT 1417
 QY 1146 GCTTTGCGAGTTGGAAGAAGTCCATACACTTTTGGATTTCTTTAAGCGGCGAGATGTC 1205
 Db 1418 GCATTTGCAATTTGGGCGCAGGATTTCTTATAATTAATTTATG---AGCTGAGTTTCTCAAGATGT 1474
 QY 1206 CAGATGCTCTCGGCTCTGATCTTCTGCTGAAGTGTGAGGAGTCTGCTGCTGCTGCTGCTGCTG 1265
 Db 1475 CAGATGTTTGGCAACTGATTTCTTGCATAAAGTGTGATGGAAGCTTTGGCAAGTCCAGGT 1534
 QY 1266 GTGCGGGGGGTGTTCTATTTGGAAGCTGGAGGAGCAAAAGCTGTGGAGGCGCATCATATCACTA 1325

Db 1535 GTGAGCGGCTGTAAATTGTGGAACTGGCAGCAAGCGGTAGAGGCAATCATAGCACTG 1594
 QY 1326 AGTGACATTTTCAAGTTCTTGCTGAGCT 1353
 Db 1595 AGTGACATATTCAACTTCTTCCITGGTT 1622

RESULT 6

AAA51354

ID AAA51354 standard; cDNA; 538 BP.

XX AC AAA51354;

DT 26-SEP-2000 (first entry)

XX Maize sucrose non-fermenting 4 protein cDNA from clore cskic.pk001.cl5.

DE Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

XX carbon catabolite repression; development; nitrogen partitioning; ss.

OS Zea mays.

XX Key Location/Qualifiers

FT CDS 347..523

FT /*tag= a

FT /product= "Partial_SNF4_protein"

FT 494..496

FT /*tag= b

FT /note= "Encodes Xaa which is not defined"

XX WO200036116-A2.

PN 22-JUN-2000.

PD 15-DEC-1999; 99WO-US023825.

PF 16-DEC-1998; 98US-0112564P.

PR (DUPO) DU FONT DE NEMOURS & CO E I.

PA Allen SM, Heppard EP, Miao G, Weng Z;

PI WPI; 2000-431594/37.

DR P-PSDB; AAY36789.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins

PT involved in carbon catabolite repression in plants and seeds; useful for

PT controlling carbon and nitrogen partitioning pathways during plant growth

PT and development.

XX Claim 18; Page 45; 48pp; English.

XX This partial cDNA encodes maize (Zea mays) sucrose non-fermenting 4

CC (SNF4) protein which is involved in carbon catabolite repression in

CC plants and seeds. The cDNA was isolated, based on similarity to SNF4

CC proteins from Arabidopsis thaliana and Saccharomyces cerevisiae, from

CC library cskic prepared from corn unpollinated developing silk 24 hours

CC after emergence. The polynucleotides are used in plants to control carbon

CC and nitrogen partitioning pathways during plant growth and development.

CC The catabolite repression proteins would facilitate studies for better

CC understanding the mechanism of catabolite repression in plants and could

CC provide genetic tools to enhance or otherwise alter the accumulation of

CC carbohydrates, lipids and proteins during plant growth and development

XX Sequence 538 BP; 146 A; 115 C; 120 G; 156 T; 0 U; 1 Other;

XX Query Watch 33.6%; Score 528; DB 3; Length 538;

XX Best Local Similarity 98.9%; Pred. No. 1.1e-137;

XX Matches 532; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 8 GGAACATTGGCGATCTCTCCATCGAAGCTGCCCACTGTATTTCAGGCTATTGCGAG 67

Db 1 GGAACATTGGCGATCTCTCCATCGAAGCTGCCCACTGTATTTCAGGCTATTGCGAG 60

QY 68 TCTGTCTCTCGGATTCACGAGTACAAATCTTGTGGACGGGAATGGCGCATGATGA 127
 Db 61 TCTGTCTCTCGGATTCACGAGTACAAATCTTGTGGACGGGAATGGCGCATGATGA 120
 QY 128 GCGTCAACCTACCATATCTCGGAGTTTGGCATAGTTAAACACACTTTTACCTTGA 187
 Db 121 GCGTCAACCTACCATATCTCGGAGTTTGGCATAGTTAAACACACTTTTACCTTGA 180
 QY 188 ATATAACCAATTAACACCTTATCACTCCCAAGCACACCTCGGAAGCAGGATGAACHTGGA 247
 Db 181 ATATAACCAATTAACACCTTATCACTCCCAAGCACACCTCGGAAGCAGGATGAACHTGGA 240
 QY 248 TGTGGATAATGAAATTTTCAACGTACGGTTTACGTCTCAGATGCCACCGTTTCAGAAGG 307
 Db 241 TGTGGATAATGAAATTTTCAACGTACGGTTTACGTCTCAGATGCCACCGTTTCAGAAGG 300
 QY 308 TACTCTGAGAGTTTCAGAGGCTGCAATPACAAATATCTAGTGTCTGTTCTGAACTCT 367
 Db 301 TACTCTGAGAGTTTCAGAGGCTGCAATPACAAATATCTAGTGTCTGTTCTGAACTCT 360
 QY 368 GAATTTGCATACATGCTATGATTTACTCCAGATTTCTGCAAGGTTATTCCTTAGCAT 427
 Db 361 GAATTTGCATACATGCTATGATTTACTCCAGATTTCTGCAAGGTTATTCCTTAGCAT 420
 QY 428 TAATTTACCTGTGAAGCAATCATTTCCATATTTCCATGAACAGGGGATTCCTGTAGCTCC 487
 Db 421 TAATTTACCTGTGAAGCAATCATTTCCATATTTCCATGAACAGGGGATTCCTGTAGCTCC 480
 QY 488 TCTCTGGGACTCAATTCAGAGTCAATTTGTTGTTGTTCTCTTAGCCCATTTGATTTCNTA 545
 Db 481 TCTCTGGGACTCAATTCAGAGTCAATTTGTTGTTGTTCTCTTAGCCCATTTGATTTCNTA 538

RESULT 7

ABL70761

ID ABL70761 standard; cDNA; 282 BP.

XX AC ABL70761;

XX 14-MAY-2002 (first entry)

DT Corn tassell-derived polynucleotide (cdps) SEQ ID NO:135.

DE Corn; corn tassell-derived polynucleotide; cdps; hybrid breeding; CDPs;

XX inheritance; characteristic; growth; development; disease resistance;

XX environmental adaptability; quality; yield; molecular marker;

XX multigene trait; plant breeding; corn tassell; gene; ss.

XX Zea mays.

OS US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999; 99US-00294093.

XX 21-APR-1998; 98US-0082567P.

XX (LALG/) LALGUDI R V.

PA (ITOL/) ITO L Y.

PA (SHER/) SHERMAN B K.

XX Lalgudi RV, Ito LY, Sherman BK;

XX WPI; 2002-163647/21.

XX Novel purified corn tassell-derived polynucleotide useful for determining

PT altered gene expression, to recover regulatory elements and to follow

PT inheritance of desirable characteristics through hybrid breeding

PT programs.

XX Claim 1; SEQ ID NO 135; 201pp; English.

XX The present sequence describes a purified corn tassel-derived
 CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
 CC selected from those given in ABL70627 to ABL76833. The cdps sequences
 CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable
 CC characteristics through hybrid breeding programs. (I) are also useful in
 CC the evaluation, and alteration of desired characteristics associated with
 CC growth and development, disease resistance, environmental adaptability,
 CC quality and yield, and as molecular markers for studying inheritance of
 CC multi-gene traits in a plant breeding program. (I) can be used to produce
 CC a tassel-specific profile of gene transcription, a transcript image, to
 CC clone regulatory elements for use in transformation vectors, to express a
 CC polypeptide, to identify, isolate or extend identical or related corn
 CC tassel nucleic acid sequences from DNA libraries, in nucleic acid
 CC hybridisation or amplification technologies, as query sequences to
 CC determine homology of known sequences, as probe for use in Southern or
 CC Northern hybridisation, and to identify the presence of and/or to
 CC determine the degree of similarity between two (or more) nucleic acid
 CC sequences
 XX
 SQ Sequence 282 BP; 86 A; 50 C; 71 G; 75 T; 0 U; 0 Other;

Query Match 16.9%; Score 266.2; DB 6; Length 282;
 Best Local Similarity 98.6%; Pred. No. 3.7e-64;
 Matches 279; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 86 CGAGTACAAATCTTTGTGAGCGGGAAATGGCGCATGATGAGCGTCAACCTACCATC 145
 DB 1 CGAGTACAAATCTTTGTGAGCGGGAAATGGCGCATGATGAGCGTCAACCTACCATC 60
 QY 146 TGGGGAGTTTGGCATAGTTTAACACACTTACTGACAGGGAATATACCAATAAACAC 205
 DB 61 TGGGGAGTTTGGCATAGTTTAACACACTTACTGACAGGGAATATACCAATAAACAC 120
 QY 206 CTTATCAAGTCCAAAGACACCTCGAAGCAGCATGATGATGATGATGATGATGATGAT 265
 DB 121 CTTATCAAGTCCAAAGACACCTCGAAGCAGCATGATGATGATGATGATGATGATGAT 180
 QY 266 TCAACGTACGTTACGTTGTGATGAGCGACCGTTTCAAGAGGTACTCTGAGGTTTCA 325
 DB 181 TCAACG-GGGGTTACGTTGTGATGAGCGACCGTTTCAAGAGGTACTCTGAGGTTTCA 239
 QY 326 GGTGCAATACAAATCTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 368
 DB 240 GGTGCAATACAAATCTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 282

RESULT 8
 ID AAA51357 standard; cDNA; 514 BP.
 XX
 AC AAA51357;
 DT 26-SEP-2000 (first entry)
 DE Wheat sucrose non-fermenting 4 protein cDNA from clone wrein.pk0143.e2.
 XX Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.
 XX
 OS Triticum aestivum.
 XX
 FH Key Location/Qualifiers
 FT CDS 19..249
 FT /*tag= a
 FT /product= "sucrose non-fermenting_4_protein"
 FT /partial
 XX
 PW WO2000036116-A2.
 XX
 PD 22-JUN-2000.

XX 15-DEC-1999; 99WO-US029825.
 PF
 XX 16-DEC-1998; 98US-0112564P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX Allen SM, Heppard EP, Miao G, Weng Z;
 PI
 XX WPI; 2000-431594/37.
 DR 2-PSDB; AAY96793.
 DR
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 PT
 XX Claim 18; Page 47; 48pp; English.
 PS
 XX This cDNA encodes a partial wheat sucrose non-fermenting 4 (SNF4) protein
 CC which is involved in carbon catabolite repression in plants and seeds.
 CC The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae. The polynucleotides
 CC are used in plants to control carbon and nitrogen partitioning pathways
 CC during plant growth and development. The catabolite repression proteins
 CC would facilitate studies for better understanding the mechanism of
 CC catabolite repression in plants and could provide genetic tools to
 CC enhance or otherwise alter the accumulation of carbohydrates, lipids and
 CC proteins during plant growth and development
 CC
 XX Sequence 514 BP; 141 A; 125 C; 104 G; 139 T; 0 U; 5 Other;

Query Match 16.4%; Score 259; DB 3; Length 514;
 Best Local Similarity 75.5%; Pred. No. 5.1e-62;
 Matches 358; Conservative 0; Mismatches 110; Indels 6; Gaps 3;
 QY 670 GACCGCAACAGCATCTAGTGCATGCCACCCCTTATAGTCTCTGAGGACATTCGAGTAA 729
 DB 5 GATCAATCAGCATCTAGTGCATGCCACCCCTTATAGTCTCTGAGGAGTATGCCATGA 64
 QY 730 AGCTTTTGAAATAGGCATTTTACAGTGCAGTATTTATTTATTCATCATCATCATCAT 789
 DB 65 AAATACCTCGAATCTGGCATTTTCTACAGTCCCAATCATCTATTCATTCGTCATCATCAT 124
 QY 790 CATTCGCCAGTATTTTGCATCTTGCATCACCTTTCTGGAATTTTGAATGTATTTGTAGAT 849
 DB 125 CGTTCCGCGAGTCTTGCATCTTGCATCCCTTTTCAAGATTTTGAATGTATTCGTAGAT 184
 QY 850 ACTTCAAAAACCTCAACTGGTAAATTTGGCTATTTCTGAAACCAACCAAGTGTCTCCATTCGCC 909
 DB 185 ACTTCAAGAACTCCACTGGTAGTTTGGCGGATTTTAAACCAACCAAGTGTCTCCATTCGCC 244
 QY 910 TGGGT--CCTGGGTTCCGAAATTTGGTGTATCTGACAGTCTGCTCATTTGGCTGTGTTG-C 966
 DB 245 TGGGGTACCTGGGGTTCGAAAAATGGGTGAACCAATGGCATCCATTTGGGTATGTTGCC 304
 QY 967 GACCTAATGCTCACTTAGCTCTCCCTTAACTATGTTGGTTCAAGCTGGAGTAAAGCTCAA 1026
 DB 305 GGCCATATACATCTCTTAACTCTGCTTAACTTGTGGGTCAAGCTGGAGTATTTATTCAA 364
 QY 1027 TACCAAT---TGTGGATGACAAACGACTCCCTGCTTGACACTTACTCTAGAGTGCATCA 1083
 DB 365 TACCCATTGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424
 QY 1084 CAGCCCTAGCAAAAGACAAGTCTTACACATGTTTCGGCTGGATGAGATGACCA 1137
 DB 425 CANTCTNGCGAAAGNAAGGCTACACCTACCGCTAGATGATGATGATGATGATGATGAT 478

RESULT 9
 ID AAA51349 standard; cDNA; 702 BP.
 XX


```
PD 25-JUL-2002.
XX
XX 18-JAN-2002; 2002WO-US001568.
XX
XX 18-JAN-2001; 2001US-0262351P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Stam L, Bachmann J, Broadus J, Kamdar KP;
XX
XX WPI; 2002-590746/63.
XX
XX P-PSDB; ABG70033.
XX
XX Identifying inhibitors of activity of proteins essential for Drosophila
XX larval viability comprises expressing in a host a protein essential for
XX larval activity and identifying compounds that inhibit or interact with
XX the protein.
XX
XX Claim 1; Page 157-158; 169pp; English.
XX
XX The invention describes a method of identifying compounds that inhibit
XX the activity of, or that interact with a protein essential for Drosophila
XX larval viability comprising expressing in a recombinant host a DNA
XX molecule to produce a protein essential for larval viability. The method
XX is useful for identifying compounds with insecticidal activity. Compounds
XX identified are useful as insecticides in crops such as maize, wheat,
XX oats, rye, sorghum, rice, barley, millet, turf, cotton, sugarcane, sugar
XX beet, oilseed rape, soybeans, vegetable crops and fruits. This sequence
XX encodes a fruit fly larval viability associated protein
XX
XX Sequence 1905 BP; 486 A; 533 C; 521 G; 365 T; 0 U; 0 Other;
XX
XX Query Match 3.2%; Score 50.6; DB 6; Length 1905;
XX Best Local Similarity 54.6%; Pred. No. 0.0024;
XX Matches 101; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
XX
XX 371 TTTCATACATGCTATGATTACTCCAGATTCTGGCAAGGTTATTGCCCTAGACATTAA 430
XX 489 TTTCACAAAGTGTATGATCTGATACCCACCTCGGCCAGTGTGTTGCTTCGACACCCA 548
XX
XX 431 TTACCTGTGACCAATCATCTCCATATCTCCATGACAGGGGATTCCTGTAGCTCTCT 490
XX 549 GCTTCTTTGAAAGAGGCTTCTACGCCCTCGTCTACACGGTGTGCGAGCGCACCGCT 608
XX
XX 491 CTGGGACTCATTCAGAGGTCATTTTGTTGCTCTCCTTAGCCCATTTGGATTTCATCTCAT 550
XX 609 CTGGGATTCGGAAGACACAGTCTGGGCGATGCTTAACCATCAGGACTTTATCAAGAT 668
XX
XX 551 ATTGC 555
XX
XX 669 CCTGC 673
XX
XX RESULT 13
XX ABL18857
XX ID ABL18857 standard; DNA; 3261 BP.
XX
XX AC ABL18857;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 8044.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 8044; 2lpp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3261 BP; 810 A; 952 C; 931 G; 568 T; 0 U; 0 Other;
XX
XX Query Match 3.2%; Score 50.6; DB 4; Length 3261;
XX Best Local Similarity 54.6%; Pred. No. 0.0031;
XX Matches 101; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
XX
XX 371 TTTCATACATGCTGATTACTCCAGATTCTGGCAAGGTTATTGCCCTAGACATTAA 430
XX 1780 TTTCACAAAGTGTATGATCTGATACCCACCTCGGCCAGTGTGTTGCTTCGACACCCA 1839
XX
XX 431 TTACCTGTGACCAATCATCTCCATATCTCCATGACAGGGGATTCCTGTAGCTCTCTCT 490
XX 1840 GCTTCTTTGAAAGAGGCTTCTACGCCCTCGTCTACACGGTGTGCGAGCGCACCGCT 1899
XX
XX 491 CTGGGACTCATTCAGAGGTCATTTTGTTGCTCTCCTTAGCCCATTTGGATTTCATCTCAT 550
XX 1900 CTGGGATTCGGAAGACACAGTCTGGGCGATGCTTAACCATCAGGACTTTATCAAGAT 1959
XX
XX 551 ATTGC 555
XX
XX 1960 CCTGC 1964
XX
XX RESULT 14
XX ABL18856
XX ID ABL18856 standard; DNA; 39651 BP.
XX
XX AC ABL18856;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 8041.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
```

XX FA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX PS Claim 1; SEQ ID NO 8041; 21pp + Sequence Listing; English.
XX PS The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL1840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 39651 BP; 11721 A; 8330 C; 8039 G; 11561 T; 0 U; 0 Other;

Query Match 3.2%; Score 50.6; DB 4; Length 39651;
Best Local Similarity 54.6%; Pred. No. 0.0095;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 371 TTTCATACATGCTATGATTACTCCAGATTCTGGCAAGGTTATTGGCCCTAGACATTAA 430
Db 36885 TTTCACAGTGCTATGATCTGTATACCACTCCGCCAAGTGTGTTCGACACCCA 36944
QY 431 TTACCTGTGAAGCAATCATTCATATTTCTCCATGAACAGGGGATTCCTGTAGCTCTCT 490
Db 36945 GCTTCTGTAAAGAAGGCTTCTAGCCCTCTGTCTACAAAGGTGTGGAGCGGACCGCT 37004
QY 491 CTGGAGCTCATTCAGAGTGCATTTGTTGGTCTCTCTTAGCCCATTTGATCTCAT 550
Db 37005 CTGGAGTTCGGAGAGCAACAGTTCGTGGGCATGCTAACCATCAGGACTTTATCAAGAT 37064
QY 551 ATTGC 555
Db 37065 CCTGC 37069

RESULT 15
ACC60984
ID ACC60984 standard; DNA; 969 BP.
XX AC ACC60984;
XX AC
XX DT 20-JUN-2003 (first entry)
XX DE Gene sequence #SEQ ID 750.
XX KW Multiprotein complex; eukaryote; drug target; diagnosis; gene; ds.
XX OS Saccharomyces cerevisiae.
XX PN EPI258494-A1.
XX PD 20-NOV-2002.
XX PF 20-DEC-2001; 2001EP-00130253.
XX PR 15-MAY-2001; 2001EP-00111774.
XX PA (CELL-) CELLZOME AG.
XX PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
PI Marzioch M, Schultz JD, Superti-Furga GD;

XX WPI; 2003-250078/25.
DR P-PSDB; ABR52942.
XX New isolated protein complexes useful for diagnosing a disease or
PT disorder, or as a target for an active agent of a pharmaceutical,
PT preferably a drug target in the treatment or prevention of disease or
PT disorder.
XX PS Disclosure; SEQ ID NO 750; 17pp + Sequence Listing; English.
XX PS The invention relates to multiprotein complexes from eukaryotes. Proteins
CC of the invention and DNA sequences encoding them are given in records
CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
CC obtainable by using a protein as a bait and isolating the set of proteins
CC which is attached thereto from cells. Such protein complexes may comprise
CC up to 30 distinct proteins. Protein complexes of the invention are useful
CC for diagnosing a disease or disorder, or as a target for an active agent
CC of a pharmaceutical, preferably a drug target in the treatment or
CC prevention of a disease or disorder. Note: The sequence data for this
CC patent is not represented in the printed specification, but is based on
CC sequence information supplied by the European Patent Office. The complete
CC document is available on CD-ROM
XX SQ Sequence 969 BP; 299 A; 173 C; 203 G; 294 T; 0 U; 0 Other;

Query Match 3.1%; Score 49.6; DB 7; Length 969;
Best Local Similarity 54.3%; Pred. No. 0.0034;
Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 361 NATATCTGAATTTGCATACATGCTATGATTACTCCAGATTCTGGCAAGGTTATTGCC 420
Db 71 AGTTTTGAATCTGAAACATCTTATGACGTGTTCCTGTCTTACCGTTTATTTGCT 130
QY 421 TAGACATTAAATTTACCTGTGAAGCAATCATTCATATTCCTCAACAGGGGATTCCTG 480
Db 131 TGGACACCTCGTTGTGTAGTGAAGAAATCACTGAATGTTCTTTGCAAAATAGCATGTCT 190
QY 481 TAGCTCTCTCTCTGGGACTCATTCAGAGTCAATTTGTTGGTCTCTTACCCCATTTGATT 540
Db 191 CTGGCCCATTTATGGGACTCCCAAGACTTCCAGGTTCTGGTGGACTTCTACTACTACAGATT 250
QY 541 TCAT 544
Db 251 TTAT 254

Search completed: July 9, 2004, 05:16:45
Job time : 723 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 03:22:12 : Search time 4369 Seconds
(without alignments)
10771.984 Million cell updates/sec

Title: US-09-857-525C-1

Perfect score: 1576

Sequence: 1 gcacgaggggaacatttgcg.....tcaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

E8T.*

1: em_estba.*

2: em_estnum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_esti.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hur.*

18: em_gss_inv.*

19: em_gss_pin.*

20: em_gss_fun.*

21: em_gss_vit.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vri.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200.6	76.2	2227	11	AY109512
2	689.6	43.8	873	14	CD434895
3	645.8	41.0	690	13	BQ163922
4	642.8	40.8	689	14	CF635900

C	5	629.4	39.9	663	14	CF011159
C	6	627.8	39.8	664	14	CF011157
	7	607.4	38.5	627	9	AW053113
	8	596.8	37.9	623	14	CF011158
	9	581.2	36.9	644	10	AW744961
C	10	580.4	36.8	601	13	BU098591
	11	575.6	36.5	616	14	CF627465
	12	574	36.4	666	14	CA211312
	13	572.8	36.3	683	14	CA262684
	14	560.4	35.6	574	14	CF011160
	15	557.4	35.4	559	10	BE123279
	16	552.8	35.1	631	14	CA253676
	17	547.4	34.7	549	13	BU499410
	18	539	34.2	738	14	CD878389
	19	537.4	34.1	595	10	BE598304
	20	520.8	33.0	534	12	BI423527
	21	520.6	33.0	680	13	CA148077
	22	518.4	32.9	712	14	CD903099
	23	516.4	32.8	737	14	CA500333
C	24	512.8	32.5	558	12	BI388981
	25	511.4	32.4	577	14	CA218557
	26	503.6	32.0	558	10	BE598914
C	27	498.2	31.6	566	14	CF624152
C	28	494	31.3	503	10	AW313249
	29	493.4	31.3	911	13	CA067477
C	30	492.6	31.3	511	9	AI987202
	31	491.8	31.2	596	14	CF052805
	32	474.6	30.1	700	13	CA124616
C	33	465	29.5	831	10	BE704511
	34	463.8	29.4	626	14	CD995954
	35	456.4	29.0	594	12	BJ450369
	36	448.4	28.5	726	14	CD902160
	37	441.4	28.0	443	9	AW056027
	38	441.2	28.0	604	12	BJ477917
	39	434.6	27.6	563	12	BG605114
	40	428	27.2	473	10	AW676933
	41	422.6	26.8	612	14	CA595438
	42	420.8	26.7	626	14	CA498139
	43	409.2	26.0	693	14	CD863501
	44	408	25.9	524	12	BG560098
	45	404.6	25.7	549	13	BU969888

ALIGNMENTS

RESULT 1

AY109512

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AY109512

Zea mays

CL367_1 mRNA sequence.

AY109512

AY109512.1

GI:21213260

HTC

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 2227)

Rainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,

Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes

Unpublished (2002)

2 (bases 1 to 2227)

Coe, E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,

these are publically available from ZmDB and may be found by BLAST

searching at MSL, maizegap.org; ZmDB, www.zmdb.lscate.edu; TIGR,

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

linear

HTC

17-OCT-2002

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

source

location/Qualifiers
1..2227
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:631725"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 76.2%; Score 1200.6; DB 11; Length 2227;
Best Local Similarity 90.8%; Pred. No. 6.3e-222; Mismatches 125; Indels 0; Gaps 0;
Matches 1239; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 8 GGAACATTTGCGGATGCTCTCTATCGAAGGCTGCCCACTGTATTTTCAAGGCTATTTGCAG 67
DB 639 GGAACATTTGCGGATGCTCTCTGTCGAAGGCTGCCCACTGTATTTTCAAGGCTATTTGCAG 698
QY 68 TCTGTCTCTGGGATTCACAGATCAAAATTTCTTGTGGACGGGAATGGCGCATGATGA 127
DB 699 CTTGTCTCCAGGATTCATGATGACAAATTTCTATGTGGACGGGAGTGGCGTCATGATGA 758
QY 128 GCGTCAACCTACCATACTCTGGGAGTTGGCATAGTTAAACACATTTACTTGAACAAGGA 187
DB 759 GCGCAACCTACTATCTGGGAGTTGGCATAGTTAAACACATTTACTTGAACAAGGA 818
QY 188 ATATAACCAATAAACCTTTATCAAGTCCAAAGCACACCTGGAAGCAGGATGAACATGGA 247
DB 819 ATTATAACCAATAAATGCTTTATTAATCCAAAGCACACCTGGAAGCAGGATGAACATGGA 878
QY 248 TGTGGATTAAGAAATTTTCAAGTACGGTACGGTACGGTACGGTACGGTACGGTACGGT 307
DB 879 TGTGGATTAAGAAATTTTCAAGTACGGTACGGTACGGTACGGTACGGTACGGTACGGT 938
QY 308 TACTCTGAGAGTTTTCAGAGGCTGCAATCAATATCTAGGTGTCGTCTTCTGAATATCT 367
DB 939 TACTCTGAGAGTTTTCAGAGGCTGCAATCAATATCTAGGTGTCGTCTTCTGAATATCT 998
QY 368 GAATTTGCATACATGCTATGATTTATCTCCAGATTCGGCAAGGTTATTCGCCCTAGACAT 427
DB 999 GAATTTGCATACATGCTATGATTTATCTCCGATTCGGCAAGGTTATTCGCCCTAGACAT 1058
QY 428 TAATTTACCTGTGAAGCAATCATTCATATTTCTCCATGACAGGAGTTCCTGTAGTCTCC 487
DB 1059 TAATTTACCTGTGAAGCAATCATTCATATTTCTCCATGACAGGAGTTCCTGTAGTCTCC 1118
QY 488 TCTCTGGAGCTATTCAGAGGCTCAATTTGTTGTTCTCTTCTAGCCCATTTGATTTTCACT 547
DB 1119 TCTCTGGAGCTATTCAGAGGCTCAATTTGTTGTTCTCTTCTAGCCCATTTGATTTTCACT 1178
QY 548 CATTTTGGGAGCTAGAACTCATGCTCGAATTCGACAGAGCAGGCTTGAACAACA 607
DB 1179 TATATTGGGAGCTAGAACTCATGCTCGAATTCGACAGAGCAGGCTTGAACAACA 1238
QY 608 CACTATATCTGCATGGAAGAGGCTAAGCGCAAACTAATGGAAGAATGATGTCAGTG 667
DB 1239 CACTATATCTGCATGGAAGAGGCTAAGCGCAAACTTGTGGAAGAATGATGTCAGTG 1298
QY 668 GCGACCCAAACAGCATCTAGTGCATGCCACCCCTTATGATTCCTTGGGACATTCAGT 727
DB 1299 GCGACCATCAGCATCTAGTGCATGCCACCCCTTATGATTCCTTGGGACATTCAGT 1358
QY 728 AAGCTTTTGCATAATGGCATTTCTACAGTGCAGGTTATTTATTCATCATCATCATGATGG 787

DB 1359 AAAGCTTTTGCTAAATGACATTTCTACAGTGCCAGTTATTTATTATCATCATCATCATGG 1418
QY 788 ATCATTCCTCCGAGTTATTCATCTTCATCATCTTTCTGAAATTTTGAATGATTGTTAG 847
DB 1419 ATCATTCCTCCGAGTTATTCATCTTCATCATCTTTCTGAAATTTTNNNNNNNNNNNN 1478
QY 848 ATACTTCAAAACCACTGCTGTAATTTGCTGATTTGCAACCAACCAAGTGTGCTCCATTC 907
DB 1479 NNN 1538
QY 908 GCTGGTTCCTGGGTTCCGAAATTTGCTGATCTGAACAGTCTCTCCATTCGCTATCTTGG 967
DB 1539 GCTGGTTCCTGGTTCGAAATTCGCTGATCTCAACAGTCTCTCATTTGGCTATCTTGG 1538
QY 968 ACCTAATGCTCATCTAGTCTCTGCTTAAATGTTGTTCAAGTGTGAGTGAAGTCAAT 1327
DB 1599 ACCTAATGCTCATCTAGTCTCTGCTTAAATGTTGTTCAAGTGTGAGTGAAGTCAAT 1658
QY 1028 ACCAATTTGCTGATGACACAGTCTCTGCTGTTGACATTTACTTACTAGAGTGCATCACAG 1087
DB 1659 ACCNN 1718
QY 1088 CCTAGCAAAAGCAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACAGGC 1147
DB 1719 CCTAGCTAAAGCAAGGCTTACACATGTTGGCTGGATGAGATGAGATGAGATGAGCA 1778
QY 1148 TTTGAGCTTGGACAGATGCAATACACATTTTGGATTTTAAAGCCGAGAGATGCA 1207
DB 1779 TTTGAGCTTGGACAGATGCAATACACATTTTGGATTTTAAAGCCGAGATGAGATGCA 1838
QY 1208 GATGCTCTCGGCTCTGATCTTCTGCTGAAGTGTGAGGACCTGCTTAATCTCTGGGT 1267
DB 1839 GATGCTCTCGGCTCTGATCTTCTGCTGAAGTGTGAGGACCTGCTTAATCTCTGGGT 1898
QY 1268 GCGGGGCTGTTCTATTTGGAAGTGGAGCAAAAGTGTGAGGAGCATCATATCACTAAG 1327
DB 1899 GCGGGGCTGTTCTATTTAGAGAGTGGAGCAAAAGTGTGAGGAGTATATATCACTGAG 1958
QY 1328 TGACATTTTCAAGTCTTCTGCTGAGCTGTGAGGAGAGGAGCA 1371
DB 1959 TGATATTTTCAAGTCTTCTGCTGAGCTGTGAGGAGAGGAGCA 2002

RESULT 2

CD434895

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Sequencing of the maize endosperm ESTs

Unpublished (2002)

Contact: Lai, Jinsheng

Dr. Joachim Messing's lab

Waksman Institute, Rutgers University

190 Frelinghuysen Rd., Piscataway, NJ 08854, USA

Tel: 732-445-3801

Fax: 732-445-5735

Email: jlai@waksman.rutgers.edu

Seq primer: T3.

Location/Qualifiers

1..873

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="W22"


```

QY 1194 GGCAGAGATCCAGATGCTGCTCGGTCTCATCTTGTCTGAGTGTGAGCGCATG 1253
Db |||||
QY 1254 GCTAATCTCGGGTGGCGGGGTGTTCTATTGTGAAGCTGGGAGCAAACTGTGTGAGGGC 1313
Db |||||
QY 1314 ATCATATCAGTATGATGATTTTCAAGTCTTCTGAGCTTGTGAGGAGAGAGCATG 1373
Db |||||
QY 1374 GCGAGCTTTTATTAACAGTAGCAACCTCTGGGCTTTGGTCTCTGGGATATTCGC 1433
Db |||||
QY 1434 TATGCTAGGAGCTCAGGTTTCATGCTGTAAAGTCTCTTTTGTGTTTTCGTTTTT 1493
Db |||||
QY 1494 ACCATTGGCCCTCCGTTGTATACAAACATACATCAGGT 1534
Db |||||
QY 1534 ACCATTGGCCCTCCGTTGTATACAAACATACATCAGGT 1

```

RESULT 4

```

LOCUS CF635900/CF635900 689 bp mRNA linear EST 02-OCT-2003
DEFINITION zmrw00.0B10-001-h03.s3 zmrw00 Zea mays cDNA, mRNA sequence.
ACCESSION CF635900
VERSION CF635900.1 GI:37397210
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

```

```

REFERENCE 1 (bases 1 to 689)
AUTHORS Bohnert, H., Sharp, R.E., Springer, G.K., Poroyko, V., Fredrickson, M.,
        Sharp, D.G., Spollen, W.G., Ries, J., Guillen, A., Khambati, A.,
        Topinka, C., Davis, G.E., Schachtman, D., Wu, Y., and Nguyen, H.T.
        NSF Grant DBI-0211842: Functional Genomics of Root Growth and Root
        Signaling Under Drought
        Unpublished (2003)
        Contact: Hans Bohnert
        University of Illinois, Urbana-Champaign
        1201 West Gregory Drive, Urbana, IL 61801, USA
        Tel: 217-265-5475
        Fax: 217-333-5574
        Email: bohnert@life.uiuc.edu
        POLYA=Yes.
        Location/Qualifiers
        1..689
        /organism="Zea mays"
        /mol_type="mRNA"
        /db_xref="taxon:4577"
        /clone_lib="zmrw00"
        /note="Samples were collected in Robert E. Sharp's lab
        (University of Missouri-Columbia) to construct three
        normalized cDNA libraries. Dark-grown maize seedlings with
        primary roots 12-20 mm in length were transplanted to high
        (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite,
        and harvested at 5 h and 48 h after transplanting. About
        1,000 roots were used for each of the low water potential
        libraries (zmrw05 and zmrw48) while 500 roots were
        combined from each of the two time points at high water
        potential (zmrw00). Each root was divided into 4 segments
        (distances are from the junction of the root apex and root
        cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7
        mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details
        of conditions see (1) with nutrient modifications as in
        (2)). The three normalized cDNA libraries were constructed
        in the lab of Hans Bohnert (University of Illinois-UC).

```

FEATURES

```

source
1..689
Query Match 40.8%; Score 642.8; DB 14; Length 689;
Best Local Similarity 96.4%; Pred. No. 4.8e-114;
Matches 671; Conservative 0; Mismatches 17; Indels 8; Gaps 1;
ORIGIN
QY 850 ACTTCAAAATCTCAACTGGTAAATTTGCTTATTCGAACCAACCACTGCTCCATTCGCG 909
Db |||
QY 689 ACTCAAAATCTCAACTGGTAAATTTGCTTATTCGAACCAACCACTGCTCCATTCGCG 630
Db |||
QY 910 TGGGTTCTCGGGTTCGGAATAATTTGGTATCTGAACAGTCTGTCCTATGTTGCGAC 969
Db |||
QY 629 TGGGTTCTCGGGTTCGGAATAATTTGGTATCTGAACAGTCTGTCCTATGTTGCGAC 570
Db |||
QY 970 CTAATGCTCTACTTAGCTCTGCCCTTACATCTGTTGTTCAAGCTGGAGTCACTCAATAC 1029
Db |||
QY 569 CTAATGCTCTACTTAGCTCTGCCCTTACATCTGTTGTTCAAGCTGGAGTCACTCAATAC 510
Db |||
QY 1030 CAAATTGTGGATGACAAACGACTCCCTGCTTGACACTTACTTAGAAGTGACATCACAGCCC 1089

```

Total RNA was extracted by the 'hot Phenol' method (Plant Molecular Biology manual, DS: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10⁶; zmrw48: 4.87 x 10⁶; zmrw00: 3 x 10⁶. The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by PCR. Plasmid DNA from the primary libraries then was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites that flank the cloned cDNA inserts. The purified PCR products, representing the entire cDNA population cloned in each library, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30°C. Non-hybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B. The total number of clones with insert was: zmrw05: 2.0x10⁷; zmrw48: 4.2x10⁷; zmrw00: 1.1x10⁷. The background of empty clones was less than 2%. Insert size, determined by PCR of the entire library, ranged from 0.5kb to 2.5kb. (1) Sharp R.E; Silk W.K; Haio T.C. Growth of the Maize Primary Root at Low Water Potentials I. Spatial Distribution of Expansive Growth. Plant Physiology (Rockville). 87(1). 1988. 50-57. (2) Spollen W.G; LeNoble M.E; Samuels T.D; Bernstein N; Sharp R.E. Abscissic acid accumulation maintains maize primary root elongation at low water potentials by restricting ethylene production. Plant Physiology (Rockville). 122(3). March, 2000. 967-976.

TAG_TISSUE=Root_segment_3
TAG_SEQ=TCGCA"


```

/organism="Zea mays"
/mol_type="mRNA"
/db_xref="F2"
/db_xref="taxon:4577"
/clone="QB16f01"
/tissue_type="pollen"
/clone_lib="QB0"

ORIGIN
Query Match      39.8%; Score 627.8; DB 14; Length 664;
Best Local Similarity 99.4%; Pred. No. 3.8e-111;
Matches 651; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 905 TCGCTGGTTCCTGGGTTCCGAAATTTGGTATCTGAACAGTGTCTGCTGCTATGTT 964
Db    |||
QY 664 TCGCTGGTTCCTGGGTTCCGAAATTTGGTATCTGAACAGTGTCTGCTGCTATGTT 605
Db    |||
QY 965 GCGACCTAATGCCCTCACTTAGCTCTGCCCTTAAACATGTTGGTCAAGCTGGAGTAAGCTC 1024
Db    |||
QY 604 CGGACCTAATGGATCACTTAGCTCTGCCCTTAAACATGTTGGTCAAGCTGGAGTAAGCTC 545
Db    |||
QY 1025 AATACCAATTTGGATGACAAAGACTCCCTGCTTGACACTTACTCTAGAAAGTGCATCAC 1084
Db    |||
QY 544 AATACCAATTTGGATGACAAAGACTCCCTGCTTGACACTTACTCTAGAAAGTGCATCAC 485
Db    |||
QY 1085 AGCCTAGCAAAAGAGAGAGTCTACACATGTTTCCGCTGGATGAGATGACCATTCACCA 1144
Db    |||
QY 484 AGCCTAGCAAAAGAGAGAGTCTACACATGTTTCCGCTGGATGAGATGACCATTCACCA 425
Db    |||
QY 1145 GGCCTTCAGCTTGCAAGAGTGCATATACACCTTTTGGATTCTTAAAGCCCAAGAGATG 1204
Db    |||
QY 424 GGCCTTCAGCTTGCAAGAGTGCATATACACCTTTTGGATTCTTAAAGCCCAAGAGATG 365
Db    |||
QY 1205 CCAGATGTCCTCGGCTCTGATCCCTTTGCTGAAGGTGATGAGGAGTGTGCTAATCTCGG 1264
Db    |||
QY 364 CCAGATGTCCTCGGCTCTGATCCCTTTGCTGAAGGTGATGAGGAGTGTGCTAATCTCGG 305
Db    |||
QY 1265 GGTGCGCGGGTGTTCATCTGAGAGTGGGAGCAAGCGTGTGGGGGCATCATATCACT 1324
Db    |||
QY 304 GGTGCGCGGGTGTTCATCTGAGAGTGGGAGCAAGCGTGTGGGGGCATCATATCACT 245
Db    |||
QY 1325 AAGTGACATTTTCAAGTTCCTGCTGAGCTTGTGAGGAGAGAGGAGTGTGGGGAGTGTTC 1384
Db    |||
QY 244 AAGTGACATTTTCAAGTTCCTGCTGAGCTTGTGAGGAGAGAGGAGTGTGGGGAGTGTTC 186
Db    |||
QY 1385 TTAAAGTAGAGACCTCTCGGGCTTTGGGCTCTGCGGATTAATTTGGCTATGCTCAGGG 1444
Db    |||
QY 185 TTAAAGTAGAGACCTCTCGGGCTTTGGGCTCTGCGGATTAATTTGGCTATGCTCAGGG 126
Db    |||
QY 1445 ACTCAGGTTTCATGCTGTAAAGTTCCTTTTGTGTTTGTCTTTTACCAATTTGGC 1503
Db    |||
QY 125 ACTCAGGTTTCATGCTGTAAAGTTCCTTTTGTGTTTGTCTTTTACCAATTTGGC 66
Db    |||
QY 1504 CCTCCCTGTTTAAACAAACATACATCAGGTGATTTTATCTATGGAATTTTC 1558
Db    |||
QY 65 CCTCCCTGTTTAAACAAACATACATCAGGTGATTTTATCTATGGAATTTTC 11
Db    |||

RESULT 7
AW053113
LOCUS 614032D06.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW053113
VERSION AW053113.1 GI:5915472
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD
c-ade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 627)
AUTHORS Walbot,V.

```

```

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 723 8221
Email: walbot@stanford.edu
Plate: 614032 row: D column: 06.

FEATURES
source 1..627
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="F23"
/db_xref="taxon:4577"
/tissue_type="root"
/lab_stage="3-4 days old"
/lab_host="XLOLR"
/clone_lib="614 - root cDNA library from Walbot Lab"
/notes="Organ: root; Vector: pBlueScriptII SK+; Site 1:
EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"

ORIGIN
Query Match      38.5%; Score 607.4; DB 9; Length 627;
Best Local Similarity 99.7%; Pred. No. 3.4e-107;
Matches 619; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 933 GGTGATCTGAACAGTCTGCTCAATTCGCTATGTTGGAGCCTTAATGCTCTACTAGCTCTGCC 992
Db    |||
QY 7 GGTGATCTGAACAGTCTGCTCAATTCGCTATGTTGGAGCCTTAATGCTCTACTAGCTCTGCC 66
Db    |||
QY 993 CTTAACATGTTGGTTCGAAGTGAAGTCAATCAATTCGAATGTCAGCAAGATGCCAAT 1052
Db    |||
QY 67 CTTAACATGTTGGTTCGAAGTGAAGTCAATCAATTCGAATGTCAGCAAGATGCCAAT 126
Db    |||
QY 1053 CTGCTTGACATTAATCTAGAGTGAATGACATGACAGCCCTAGCAAAAGAGAGTCTACACA 1112
Db    |||
QY 127 CTGCTTGACATTAATCTAGAGTGAATGACATGACAGCCCTAGCAAAAGAGAGTCTACACA 186
Db    |||
QY 1113 CATGCTCGCTGGATGAGATGACCAATTCACAGGCTTTGCAAGTGTGCAAGATGCCAAT 1172
Db    |||
QY 187 CATGCTCGCTGGATGAGATGACCAATTCACAGGCTTTGCAAGTGTGCAAGATGCCAAT 246
Db    |||
QY 1173 ACACCTTTGGATCTTTAAAGCCAGAGATGCGCAGATGTCCTCCGGCTCTGATCTTTG 1232
Db    |||
QY 247 ACACCTTTGGATCTTTAAAGCCAGAGATGCGCAGATGTCCTCCGGCTCTGATCTTTG 306
Db    |||
QY 1233 CTGAAGTGTGAGGAGGAGTGTGCTTAATCTCGGGGTGCGGGGTGTTTCATTTGGAAGCT 1292
Db    |||
QY 307 CTGAAGTGTGAGGAGGAGTGTGCTTAATCTCGGGGTGCGGGGTGTTTCATTTGGAAGCT 366
Db    |||
QY 1293 GGGAGCAAAAGTGTGAGGAGGAGTGTGCTTAATCTCGGGGTGCGGGGTGTTTCATTTGGAAGCT 1352
Db    |||
QY 367 GGGAGCAAAAGTGTGAGGAGGAGTGTGCTTAATCTCGGGGTGCGGGGTGTTTCATTTGGAAGCT 426
Db    |||
QY 1353 TTGTGAGGAGGAGGAGGAGTGTGCTTAATCTCGGGGTGCGGGGTGTTTCATTTGGAAGCT 1412
Db    |||
QY 427 TTGTGAGGAGGAGGAGGAGTGTGCTTAATCTCGGGGTGCGGGGTGTTTCATTTGGAAGCT 486
Db    |||
QY 1413 GGGTCTTTGGAGTAAATTTGGTATCTGTCAGGAGTCTAGGGTTCATGCTCTGAAAGTTC 1472
Db    |||
QY 487 GGGTCTTTGGAGTAAATTTGGTATCTGTCAGGAGTCTAGGGTTCATGCTCTGAAAGTTC 546
Db    |||
QY 1473 -TTTTTTTTTGTGTTTTTACCATTGGCCCTCCCGTGTGTTTAAACAAACATACATCA 1531
Db    |||
QY 547 TTTTTTTTTTGTGTTTTTACCATTGGCCCTCCCGTGTGTTTAAACAAACATACATCA 606
Db    |||
QY 1532 GGTGATTTTATCTATTGGAT 1552
Db    |||
QY 607 GGTGATTTTATCTATTGGAT 627
Db    |||

```

```

RESULT 8
CF011158
LOCUS      CF011158                623 bp    mRNA    linear    EST 17-JUL-2003
DEFINITION OBJ6f01.xg OBJ Zea mays cDNA clone QBU6f01, mRNA sequence.
ACCESSION  CF011158
VERSION     CF011158.1  GI:32906345
KEYWORDS   EST.
SOURCE     Zea mays
           Zea mays
ORGANISM   Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade; Panicoideae; Andropogoneae; Zea.
           1 (bases 1 to 623)
REFERENCE   Genoplante, a major partnership french program in p-ant genomics
           Genoplante, (2003)
           Unpublished (2003)
           Contact: Genoplante
           Genoplante
           93, rue Henri Rochefort 91025 EVRY CEDEX France
           Tel: 33 1 69 47 54 00
           Fax: 33 1 69 47 54 10
           This sequence has been generated in the framework of the french
           plant genomics programme 'Genoplante' (http://www.genoplante.com)
           and http://genoplante-info.infobiogen.fr.
FEATURES   Location/Qualifiers
             1..623
             /organism="Zea mays"
             /mol_type="mRNA"
             /cultivar="P2"
             /db_xref="taxon:4577"
             /clone="QBU6f01"
             /tissue_type="pollen"
             /clone_lib="QBU"
ORIGIN
Query Match      37.9%; Score 596.8; DB 14; Length 623;
Best Local Similarity 99.4%; Pred. No. 3.8e-105;
Matches 620; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 905 TCCGCTGGGTTCTCGGTTCCGAAATTTGGTATCTGACAGTCGTCATGGCTATGTT 964
Db 1 TCCGCTGGGTTCTCGGTTCCGAAATTTGGTATCTGACAGTCGTCATGGCTATGTT 60

QY 965 GCGACCTAATGCTCATTAGCTCTGCCCTTAACATGTTGGTTCAAGCTGGAGTAAGCTC 1024
Db 61 GCGACCTAATGCTCATTAGCTCTGCCCTTAACATGTTGGTTCAAGCTGGAGTAAGCTC 120

QY 1025 AATACCAATTTGGATGACACGACCTCCCTGCTTGACACTTACTCTAGAAGTGACATCAC 1084
Db 121 AATACCAATTTGGATGACACGACCTCCCTGCTTGACACTTACTCTAGAAGTGACATCAC 180

QY 1085 AGCCCTAGCAAAAGACAAGGCTTCACACATGTTGGCTGGATGAGATGACCATTCACCA 1144
Db 181 AGCCCTAGCAAAAGACAAGGCTTCACACATGTTGGCTGGATGAGATGACCATTCACCA 240

QY 1145 GCTTTGACGCTTGACACAGATGCCAATACACTTTTGGATTCTTTAAACGGCCAGAGATG 1204
Db 241 GCTTTGACGCTTGACACAGATGCCAATACACTTTTGGATTCTTTAAACGGCCAGAGATG 300

QY 1205 CCAGATGCGCTCCGGTCTGATCTTTGCTGAAGTGATGGACGCTGGCTAACTCGG 1264
Db 301 CCAGATGCGCTCCGGTCTGATCTTTGCTGAAGTGATGGACGCTGGCTAACTCGG 360

QY 1265 GGTGCGCGGGTGTTCAATGTGGAGCTCGGAGCAAACTGTGGGGCATCATATCACT 1324
Db 361 GGTGCGCGGGTGTTCAATGTGGAGCTCGGAGCAAACTGTGGGGCATCATATCACT 420

QY 1325 AAGTGACATTTTCAAGTCTTCTGAGCTTGTGAGGAGAGAGCATGCGGAGTTTTT 1384
Db 421 AAGTGACATTTTCAAGTCTTCTGAGCTTGTGAGGAGAGAGCATGCGGAGTTTTT 479

```

```

QY 1385 TTAAACAGTAGCAACCCCTCTGGGCTTTGGGCTCTTGGGATAATTGGCTATCGTCAGGG 1444
Db 480 TTAAACAGTAGCAACCCCTCTGGGCTTTGGGCTCTTGGGATAATTGGCTATCGTCAGGG 539

QY 1445 ACTCAGGTTTCATGCTGCTGTAAGTTCC-TTTTTTTTGGTTCGTTTTTACCAATTGGC 1503
Db 540 ACTCAGGTTTCATGCTGCTGTAAGTTCC-TTTTTTTTGGTTCGTTTTTACCAATTGGC 599

QY 1504 CCTCCCGTCTGTAACAAACATAC 1527
Db 600 CCTCCCGTCTGTAACAAACATAC 623

RESULT 9
AW744961
LOCUS      AW744961                644 bp    mRNA    linear    EST 19-JUL-2000
DEFINITION LG1_385_D10.b1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
           sequence.
ACCESSION  AW744961
VERSION     AW744961.1  GI:7658699
KEYWORDS   EST.
SOURCE     Sorghum bicolor (sorghum)
           Sorghum bicolor
           Sorghum bicolor
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade; Panicoideae; Andropogoneae; Sorghum.
           1 (bases 1 to 644)
REFERENCE   Cordonnier-Pratt M.-M., Gingle A., Marsala C. and Pratt L.H.
           An EST database from Sorghum: light-grown seedlings
           Unpublished (2000)
           Contact: Cordonnier-Pratt MM
           Laboratory for Genomics and Bioinformatics
           The University of Georgia, Department of Plant Biology
           Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
           Tel: 706 542 1860
           Fax: 706 583 0210
           Email: mmpratt@uga.edu
           Sequences have been trimmed to exclude PolyA, vector and regions
           below Phred quality 16. The threshold for highest quality sequence
           is 20.
Seq primer: JEN REV
High quality sequence stop: 621
POLYA=No.
FEATURES   Location/Qualifiers
             1..644
             /organism="Sorghum bicolor"
             /mol_type="mRNA"
             /db_xref="taxon:4558"
             /note="Organ: 10- to 14-day-old light-grown (greenhouse)
             seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2:
             EcoRI; The library was made from poly-A RNA in the cloning
             vector lambda ZAP II. Clones to be sequenced were
             prepared by mass excision."
ORIGIN
Query Match      36.9%; Score 581.2; DB 10; Length 644;
Best Local Similarity 94.1%; Pred. No. 3.8e-102;
Matches 604; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 8 GGAAACATTTGCCGATGCTCTCTATCGAAGCTGCCACCTGTAATTCAGCTATTTCAG 67
Db 3 GGAAACATTTGCCGATGCTCTCTATCGAAGCTGCCACCTGTAATTCAGCTATTTCAG 62

QY 68 TCTGTCTCTCGGATTCAGAGTCAAAATTCCTTGTGGACGGGATGGCGCATATGA 127
Db 63 CTGTCTCTCGGATTCAGAGTCAAAATTCCTTGTGGATGGGAGTGGCGCATATGA 122

QY 128 GCGTCAACCTACCATATCTGGGGAGTTTGGCATAGTTTAAACACTTTACTTGGACAAGGA 187
Db 123 GCGCAACCTACTATCTGGGGAGTTTGGATAGTTTAAACACTTATCTTGGACAAGGA 182

QY 188 ATATACCAAAATAAACACCTTTATCAAGTCCAAAGCACCTCGGAAGCAGGATGAACATGA 247

```

```

Db 183 ATTTACCAACTAAAGCCCTTTATAGTCCAGACACCTGGAAGCAGGATGAACATGGA 242
Qy 248 TGTGATAATGAAATTTTCAACGTACGGTTACGTTGTCAGATGCCACCGTTTCAGAGG 307
Db 243 TGTGATAATGAAATTTTCAACGTACGGTTACGTTGTCAGATGCCACCGTTTCAGAGG 302
Qy 308 TACTCTGAGAGTTTCAGAGGTCGCAATACAAATATCTAGGTCGTTGTTCTGGAATATCT 367
Db 303 TACTCCGAGAGTTTCAGAGGTCGCAATACAAATCTTAGGTCGCGGTTTCGGAATATCT 362
Qy 368 GAATTTGATACATCTGATGATTTACTCCAGATTTCCGCAAGTTATTTGCCCTAGACAT 427
Db 363 GAATTTGATACATCTGATGATTTACTCCGCAAGTTTCGCGCAAGTTATTTGCCCTAGACAT 422
Qy 428 TAATTTACCTGTGAAGCAATCATTCATCTCCATGAAACAGGAGATTCCTGTAGCTCC 487
Db 423 TAATTTACCTGTGAAGCAATCATTCATCTCCATGAAACAGGAGATTCCTGTAGCTCC 482
Qy 488 TCTCTGGGACTCATCTGAGAGTCATTTGTTGGTCTCCTTAGCCCATTTGGATTTCACT 547
Db 483 TCTCTGGGACTCATCTGAGAGTCATTTGTTGGTCTCCTTAGCCCATTTGGATTTCACT 542
Qy 548 CATATATCTGCATGGAAGAGGCTTAGCGGCAAACTAATGG 649
Db 543 TATATTGGGGAGCTAGAACTCATGCTCGAATTTGACAGAGGAGCTTGAACACA 602
Qy 608 CATATATCTGCATGGAAGAGGCTTAGCGGCAAACTAATGG 649
Db 603 TACGATATCTGCATGGAAGAGGCTTAGCGGCAAACTAATGG 644

```

```

RESULT 10
BU098591
LOCUS 946136F11.y1 946 - tassels primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.

```

```

ACCESSION BU098591
VERSION BU098591.1 GI:22546280
KEYWORDS EST.
SOURCE Zea mays

```

```

ORGANISM Zea mays

```

```

REFERENCE 1 (bases 1 to 601)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

```

```

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V

```

```

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946136 row: F column: 11.
Location/Qualifiers
1. .601
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XL0LR"
/clone_lib="946 - tassels primordium prepared by Schmidt
lab"

```

```

FEATURES
source

```

Library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

ORIGIN

```

Query Match 36.8%; Score 580.4; DB 13; Length 601;
Best Local Similarity 99.8%; Pred. No. 5.6e-102;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 8 GGAACATTTGGCGATGTCCTCTATCGAAGGCTGCCCACTGTTATTTTCAGGCTATTTGCAG 67
Db 20 GGAACATTTGGCGATGTCCTCTATCGAAGGCTGCCCACTGTTATTTTCAGGCTATTTGCAG 79
Qy 68 TCTGTCTCTCGGATTCACGAGTACAAATCTTTGTGACGGGGAATCGCGCATGATGA 127
Db 80 TCTGTCTCTCGGATTCACGAGTACAAATCTTTGTGACGGGGAATCGCGCATGATGA 139
Qy 128 GCCTCAACCTTACCATACTCTGGGGAGTTTGGCATAGTTAAACACATTTTACTTGACAAGGA 187
Db 140 GCCTCAACCTTACCATACTCTGGGGAGTTTGGCATAGTTAAACACATTTTACTTGACAAGGA 199
Qy 188 ATATAACCAATTAACACCTTTATCAAGTCCAAAGCACAACCTGGAAGCAGGATGAACATGA 247
Db 200 ATATAACCAATTAACACCTTTATCAAGTCCAAAGCACAACCTGGAAGCAGGATGAACATGA 259
Qy 248 TGTGATTAATCAAAATTTTCAACGTACCGTTTACGTTGTTCAGATGCAACGTTTCAGAAAG 307
Db 260 TGTGATTAATCAAAATTTTCAACGTACCGTTTACGTTGTTCAGATGCAACGTTTCAGAAAG 319
Qy 308 TACTCTGAGAGTTTCAGAGGTCGCAATACAAATATCTAGGTCGTGTTTTCGAATATCT 367
Db 320 TACTCTGAGAGTTTCAGAGGTCGCAATACAAATATCTAGGTCGTGTTTTCGAATATCT 379
Qy 368 GAATTTGCATACATGCTATGATTTTACTCCAGATTTCTGCAAGGTTATTTGCCCTAGACAT 427
Db 380 GAATTTGCATACATGCTATGATTTTACTCCAGATTTCTGCAAGGTTATTTGCCCTAGACAT 439
Qy 428 TAAATTTACTCTGTGAAGCAATCATTCATATTTCTCCATGAACAGAGGGAATTCCTGTAGCTCC 487
Db 440 TAAATTTACTCTGTGAAGCAATCATTCATATTTCTCCATGAACAGAGGGAATTCCTGTAGCTCC 495
Qy 488 TCTCTGGGACTCATCTCAGAGTCATATTTGTTGGTCTCCTTAGCCCATTTGGAATTTCACT 547
Db 500 TCTCTGGGACTCATCTCAGAGTCATATTTGTTGGTCTCCTTAGCCCATTTGGAATTTCACT 559
Qy 548 CATATTTGGGAGGCTAGAACTCATGCTCGCAACTTTGACAGA 589
Db 560 CATATTTGGGAGGCTAGAACTCATGCTCGCAACTTTGACAGA 601

```

RESULT 11

```

CF627465/c
LOCUS CF627465.1 616 bp mRNA linear EST 02-OCT-2003
DEFINITION zmrws05.0B20-011-f01.s4 zmrws05 Zea mays cDNA, mRNA sequence.
ACCESSION CF627465
VERSION CF627465.1 GI:37381198
KEYWORDS EST.
SOURCE Zea mays

```

ORGANISM

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

```

```

REFERENCE 1 (bases 1 to 616)
AUTHORS Bohnert,H., Sharp,R.E., Springer,G.K., Poroyko,V., Fredrickson,M.,

```

```

Sharp,L.G., Spollen,W.G., Rieg,J., Guillen,A., Khambati,A.,
Topinka,C., Davis,G.E., Schachtman,D., Wu,Y. and Nguyen,H.T.
NSF Grant DBI-0211842; Functional Genomics of Root Growth and Root
Signaling Under Drought

```

```

JOURNAL Unpublished (2003)
COMMENT Contact: Hans Bohnert

```

```

University of Illinois, Urbana-Champaign
1201 West Gregory Drive, Urbana, IL 61801, USA
Tel: 217-265-5475
Fax: 217-333-5574

```

Email: bohnert@life.uiuc.edu
POLYA=Yes.

FEATURES

source
1. .616
Location/Qualifiers
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="zrrws05"

/notes="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zrrws05 and zrrws48) while 500 roots were combined from each of the two time points at high water potential (zrrws00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (root segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zrrws05: 3.37 x 10⁶; zrrws48: 4.87 x 10⁶; zrrws00: 3 x 10⁶. The background of empty clones was less than 1%. Inserts ranged from ~0.5 kb to >2.5 kb, as determined by PCR. Plasmid DNA from the primary libraries then was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites that flank the cloned cDNA inserts. The purified PCR products, representing the entire cDNA population cloned in each library, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 300C. Non-hybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B. The total number of clones with insert was: zrrws05: 2.0x10⁷; zrrws48: 4.2x10⁷; zrrws00: 1.1x10⁷. The background of empty clones was less than 2%. Insert size, determined by PCR of the entire library, ranged from 0.5kb to 2.5kb. (1) Sharp R E; Silk W K; Hsiao T C. Growth of the Maize Primary Root at Low Water Potentials I. Spatial Distribution of Expansive Growth. Plant Physiology (Rockville) 87(1): 1986. 50-57. (2) Spollen W G; LeNoble M E; Samuels T D; Bernstein N; Sharp R E. Abscisic acid accumulation maintains maize primary

root elongation at low water potentials by restricting ethylene production. Plant Physiology (Rockville) 122(3): March, 2000. 967-976.
TAG_TISSUE=Root_segment_4
TAG_SEQ=TCGGA"

ORIGIN

Query Match 36.5%; Score 575.6; DB 14; Length 616;
Best Local Similarity 96.5%; Pred. No. 4.6e-101;
Matches 602; Conservative 0; Mismatches 14; Indels 8; Gaps 1;
Qy 935 TGATCTGAACAGTCGTCAATGGCTATGTGCGACCAATTAAGCCCTACATTAGCTGSCCT 994
Db 616 TGATCTGAACAGTCGTCAATGGCTATGTGCGACCAATTAAGCCCTACATTAGCTGSCCT 557
Qy 995 TACATCTGGTTCAGCTGAGTAAAGCTCAATACCAATTCGATGATGACACGACTCCCT 1054
Db 556 TACATCTGGTTCAGCTGAGTAAAGCTCAATACCAATTCGATGATGACACGACTCCCT 497
Qy 1055 GCTTGACACTTACTCTAGAGTGCATCAGCCCTAGCAAAAGACAGGCTCTACACACA 1114
Db 496 GCTTGACACTTACTCTAGAGTGCATCAGCCCTAGCAAAAGACAGGCTCTACACACA 437
Qy 1115 TGTTCGGCTGATGAGATGACCAATTCACAGGCTTTCAGCTTGGACAAAGATGCCATAC 1174
Db 436 TGTTCGGCTGATGAGATGACCAATTCACAGGCTTTCAGCTTGGACAAAGATGCCATAC 377
Qy 1175 ACCTTTTGGATCTTTAAACGCCAGAGATGCCAGATGCGCTCCGGCTGATCCTTTCCT 1234
Db 376 ACCTTTTGGATCTTTAAACGCCAGAGATGCCAGATGCGCTCCGGCTGATCCTTTCCT 317
Qy 1235 GAAGTGATGAGGAGCACTGGCTAATCTCGGGTTCGGGGGTGTTTCATTGTGAGAGCTGG 1294
Db 316 GAAGTGATGAGGAGCACTGGCTAATCTCGGGTTCGGGGGTGTTTCATTGTGAGAGCTGG 257
Qy 1295 GAGCAACAGTGTGAGGGGATCATATCACTAAGTGCATTTTCAAGTCTTCTGCTGAGCTT 1354
Db 256 GAGCAACAGTGTGAGGGGATCATATCACTAAGTGCATTTTCAAGTCTTCTGCTGAGCTT 197
Qy 1355 GTGAGGAGAGAGAGCATGCGGAGTGTTTTAAACAGTAGCAACCCCTCTGGGCTTTGG 1414
Db 196 GTGAGGAGAGAGAGCATGCGGAGTGTTTTAAACAGTAGCAACCCCTCTGGGCTTTGG 137
Qy 1415 GTCTTCGGGATTAATTCGCTATCTGAGGAGCTCAGGGTTCATGGTCTGTAAGTTCCTT 1474
Db 136 GTCTTCGGGATTAATTCGCTATCTGAGGAGCTCAGGGTTCATGGTCTGTAAGTTCCTT 85
Qy 1475 TTTTTCGTTTCGTTTTCATTCATTTTACCATTTGGCCCTCCCGTTGTTAAACAAATACATCAGGT 1534
Db 84 TTTTTCGTTTCGTTTTCATTCATTTTACCATTTGGCCCTCCCGTTGTTAAACAAATACATCAGGT 25
Qy 1535 GATTTTTCATCTATTGGATTCGTTTC 1558
Db 24 GATTTTTCATCTATTGGATTCGTTTC 1

RESULT 12

LOCUS CA211312 666 bp mRNA linear EST 25-SEP-2003
DEFINITION SCRLAD1098B08.g AD1 Saccharum officinarum cDNA clone SCRLAD1098B08
CA211312
ACCSSION CA211312
VERSION CA211312.1 GI:35255493
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
REFERENCE
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 666)
Vettore A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCRST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

COMMENT

Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 098 row: B column: 08
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .666
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCELAD1098B08"
/lab_host="DH10B"
/clone_lib="Ad1"
/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Gluconacetobacter
diazotrophicans]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucet.lad.ic.unicamp.br/public"

FEATURES

source

ORIGIN

Query Match 36.4%; Score 574; DB 14; Length 666;
Best Local Similarity 93.0%; Pred. No. 9e-101;
Matches 612; Conservative 0; Mismatches 45; Indels 1; Gaps 1;
QY 12 CATTTGGCGATGCTCTCTATCGAAGGCTGCCCTCTGTATTTTCAGGCTATTTGCGATCTG 71
DB 1 CATTTGGCGATGCTCTCTCTATCGAAGGCTGCCCTCTGTATTTTCAGGCTATTTGCGATCTG 60
QY 72 TCTCTGGGATTCACGAGTACAAATCTCTGTGACGGGGAATCGCGCATGATGAGCGT 131
DB 61 TCTCCAGGATCCACGAGTACAGATTTTGTGACGGGAGTGGCGCATGATGAGCGC 120
QY 132 CAACCTACCATATCTGGGAGTTTGGCATAGTTTAACACACTTTACTTGCACAAAGGGAATAT 191
DB 121 CAACCTACTATATCCGGGAGTTTGGTATAGTTTAACACACTTTACTTGCACAAAGGAATTT 180
QY 192 AACCAATACACACTTATCAAGTCCAGACACACTGGAAGCAGGATGACATGATGTG 251
DB 181 AACCAATACACGCTTATTAAAGTCCAAAGCAATTTGGAAGCAGGATGAAATCGATGTG 240
QY 252 GATAATGAAATTTTCAACGTACGGTTACGTTGTGAGATGACCGTTTCAGAGGACT 311
DB 241 GATATGAAATTTTCAACGTACGGTTACGTTGTGAGATGACCGTTTCAGAGGACT 300
QY 312 CTGAGATTTTCAGAGGCTGCAATACAAATATCTAGGTGTGCTGTCTGTAATATCTGAAT 371
DB 301 CCGAGAGTTTCAGAGGCTGCAATACAAATCTAGGTGTGCTGTCTGTAATATCTGAAT 360
QY 372 TTGCATACATGCTATGATTTTACTCCAGATTTCTGCGAGGTTATTCCTAGATTAAT 431
DB 361 TTGCATACATGCTATGATTTTACTCCGAGTTTCAGCGAGGTTATTCCTAGATTAAT 420
QY 432 TTACTGTGAGGCAATCATTCATATTTCTCCATGAAACAGGGGATTCCTGTAGCTTCCTC 491
DB 421 TTACTGTGAGGCAATCTTTTCATATTTCTATGACAGGGGATTCCTGTAGCTTCCTC 480
QY 492 TGGGACTCATTCAGAGGTCMAATTTGTTGGTCTCCTTAGCCCATTTGGAATTCATCTCAT 551
DB 481 TGGGACTCATTCAGAGGTCMAATTTGTTGGTCTCCTTAGCCCATTTGGAATTCATCTCAT 540

QY 552 TTGCGGAGCTAGAAACTCATGGCTCGAACTTGACAGAGAGCAGCTTGAAACACACACT 611
DB 541 TTGCGGAGCTAGAAACTCATGGCTCGAACTTGACAGAGAGCAGCTTGAAACACACACT 600
QY 612 ATATCTGCATGGAAGAGGCTTAAGCGCAAC-TAATGGAAGAAATGATAGTCACTGG 668
DB 601 ATATCTGCATGGAAGAGGCTTAAGCGCAACCTTTATCGAAGAAATGATAGTCAATTG 658

RESULT 13

CA262684

LOCUS

DEFINITION

CA262684

VERSION

CA262684.1

KEYWORDS

EST.

SOURCE

ORGANISM

Saccharum officinarum

Saccharum officinarum

REFERENCE

1 (Bases 1 to 683)

AUTHORS

Vettore, A.L., da Silva, P.R., Kemper, E.L. and Arruda, P.

TITLES

The libraries that made SUCEST

JOURNAL

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

COMMENT

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 024 row: A column: 07

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. .683

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCPILB2024A07"

/lab_host="DH10B"

/clone_lib="LB2"

/note="Organ: Lateral buds from plants adult plants

growing in greenhouse; Vector: pSport1; Site 1: SalI;

Site 2: NotI; An unidirectional cDNA library generated

from [lateral buds from plants adult plants growing in

greenhouse]. cDNA was prepared from polyA+ mRNA using

SuperScript Plasmid System Kit (Invitrogen). The

double-strand cDNAs were fractionated in a sepharose

CL-2B 40cm-columns and fragments sizing between 0.8 and

1.5 Kb were directionally cloned into the vector. Details

of each source of RNA and library construction can be

obtained at http://sucet.lad.ic.unicamp.br/public"

ORIGIN

Query Match 36.3%; Score 572.8; DB 14; Length 683;
Best Local Similarity 92.7%; Pred. No. 1.5e-100;
Matches 612; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
QY 8 GGAACATTTGCGATGCTCTCTATCGAAGGCTGCCCTCTGTATTTTCAGGCTATTTCCAG 67
DB 25 GGAACATTTGCGATGCTCTCTCTGTGTA-GTTGCCCTCTGTATTTTCAGGCTATTTCCAG 83
QY 68 TCTGTCCTCTGGATTCACGAGTACAAATCTTTGTGACGGGAATGCGGATGATGA 127
DB 84 CCGTCTCCAGGGATTCACGAGTACAAAGCCCTTTTGTGACGGGAGTGGCGGATGATGA 143
QY 128 GCGTCAACCTACCATATCTCGGGAGTTTGGCATAGTTTAAACACACTTTTACTTTGACAGGGA 187

Db 144 GCGCAACCTACTATATCCGGGGAGTTTGGTATAGTTAAACACATGTACTTGAACAAGAGA 203

Qy 188 ATATAACCAATAAACAACCTTATCAAGTCCAGACACACCTGGAGGAGGATGACATGGA 247

Db 204 ATTAAACCAATAAACAACCTTATTAAGTCCAGACCAATTTGAAGAGGAGGATGACATGGA 263

Qy 248 TGTGGAATGAATAATTTCAACGTACGGTTACGTTGTGATGGCCACCGTTTCAGAAAGG 307

Db 264 TGTGGAATGAATAATTTCAACGTACGGTTACGTTGTGATGGCCACCGTTTCAGAAAGG 323

Qy 308 TACTCTGAGATTTCAGAGGTCGAATACAAATATCTAGGTGTCGTTTCTGAATATCT 367

Db 324 TACTCCGAGAGTTTCAGAGGTCGAATACAAATCTCTAGGTGTCGTTTCTGAATATCT 383

Qy 368 GAATTTGCATACATGCTATGATTTACTCCAGATTTCTGGCAGGTTTATGGCCCTAGACNT 427

Db 384 GAATTTGCATACATGCTATGATTTACTCCGAGTTTCAGGCAAGGTTTATGGCTCGGACNT 443

Qy 428 TAATTTACCTGTGAAGCAATCAATCCCAATTTCTCCATGAACAGGGGATTTCTGTAGCTCC 487

Db 444 TAATTTACCTGTGAAGCAATCTTTTCAATTTCTACATGAACAGGGGATTTCTGTAGCTCC 503

Qy 488 TCTCTGAGACTCATTCAGAGTCAATTTGTTGGTCTCTTAGCCCATTTGGAATTTCAACT 547

Db 504 TCTCTGAGACTCATTCAGAGTCAATTTGTTGGTCTCTTAGCCCATTTGGAATTTCAACT 563

Qy 548 CATATTGCGGGAGCTAGAACTCATGGCTCGAACTTCACAGACAGAGCAGCTTGAACACA 607

Db 564 TATATTGCGGGAGCTAGAACTCATGGCTCGAACTTCACAGACAGAGCAGCTTGAACACA 623

Qy 608 CACTATATCTGCATGGAAGAGGCTAAGCGGCAAACTTAATGGAAGAAATGATAGTCAGTG 667

Db 624 CACTATATCTGCATGGAAGAGGCTAAGCGGCAAACTTAATGGAAGAAATGATAGTCAGTG 683

RESULT 14

CF011160

LOCUS QBj6f03.xg QBj Zea mays cDNA clone QBj6f03, mRNA sequence. EST 17-JUL-2003

DEFINITION

ACCESSION CF011160

VERSION CF011160.1 GI:32906347

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 574)

Genoplatte.

Genoplatte, a major partnership french program in plant genomics

Unpublished (2003)

Contact: Genoplatte

Genoplatte

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>) and <http://genoplatte-info.infnbiogen.fr>.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .574

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="P2"

/db_xref="taxon:4577"

/clone="QBj6f03"

/tissue_type="pollen"

/clone_lib="QBj"

ORIGIN

Query Watch 35.6%; Score 560.4; DB 14; Length 574;

Best Local Similarity 99.5%; Pred. No. 4.1e-98;

Matches 572; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 905 TCCGCTGGGTTCCCTGGGTTCCGAAATTTGGTGAATGTAACAGTCTGTCATTTGGCTATGTT 964

Db 1 TCCGCTGGGTTCCCTGGGTTCCGAAATTTGGTGAATGTAACAGTCTGTCATTTGGCTATGTT 60

Qy 965 GCGACCTAATAGCTTCACTTAGCTCTGCGCTTTAAACATGTTGGTTCAAGCTGGAGTAAGCTC 1024

Db 61 GCGACCTAATAGCTTCACTTAGCTCTGCGCTTTAAACATGTTGGTTCAAGCTGGAGTAAGCTC 120

Qy 1025 AATACCAATTTGGTGAATGACAAAGACTCCCTGCTTGAACATTTACTTCTAGAGTGAATCAC 1084

Db 121 AATACCAATTTGGTGAATGACAAAGACTCCCTGCTTGAACATTTACTTCTAGAGTGAATCAC 180

Qy 1085 AGCCTAGCAAAAGACAGAGCTTACACACATGTTGGCTGGATGAGATGACCATTCACCA 1144

Db 181 AGCCTAGCAAAAGACAGAGCTTACACACATGTTGGCTGGATGAGATGACCATTCACCA 240

Qy 1145 GGCTTTGAGCTTGGCAAGATGCAATACACCTTTTGGATTTCTTAACGGCCAGAGATG 1204

Db 241 GGCTTTGAGCTTGGCAAGATGCAATACACCTTTTGGATTTCTTAACGGCCAGAGATG 300

Qy 1205 CCAGATGCTCCCTGGTCTGATCTTTGCTGAAAGTGAAGAGGAGCTGGCTAATCTCTGG 1264

Db 301 CCAGATGCTCCCTGGTCTGATCTTTGCTGAAAGTGAAGAGGAGCTGGCTAATCTCTGG 360

Qy 1265 GGTGCGGGGGTGTTCATTTGGAGAGCTGGGAGCAAACTGTGGAGGAGCATCATATCACT 1324

Db 361 GGTGCGGGGGTGTTCATTTGGAGAGCTGGGAGCAAACTGTGGAGGAGCATCATATCACT 420

Qy 1325 AAGTGACATTTTCAAGTTCTTGTGAGCTTGTGAGAGAGAGGAGCATGGCGGAGTTTTT 1384

Db 421 AAGTGACATTTTCAAGTTCTTGTGAGCTTGTGAGAGAGAGGAGCATGGCGGAG-TTTT 479

Qy 1385 TTAAACAGTAGCAACCTCTGGGCTTTGGGCTTTCGGGTAATTTGGCTATCGTCAGGG 1444

Db 480 TTAAACAGTAGCAACCTCTGGGCTTTGGGCTTTCGGGTAATTTGGCTATCGTCAGGG 539

Qy 1445 ACTCAGGTTTCATGCTGTGAAGCTTCTCTTTT 1479

Db 540 ACTCAGGTTTCATGCTGTGAAGCTTCTCTTTT 574

RESULT 15

BE123279

LOCUS

DEFINITION BE123279 559 bp mRNA linear EST 14-JUN-2003

{SK} Zea mays cDNA, mRNA sequence.

ACCESSION BE123279

VERSION BE123279.1 GI:8516554

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 559)

Walbot V

Maize ESTs from various cDNA libraries sequenced at Stanford University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 945040 row: D column: 06.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .559

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="W23"

/db_xref="taxon:4577"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stages="fully-grown"
/lab_host="DH128"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/note="organ: tassel, kernal, silk, husk, root, leaf;
vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
-issue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

ORIGIN

Query Match 35.4%; Score 557.4; DB 10; Length 559;
Best Local Similarity 99.8%; Pred. No. 1.6e-97;
Matches 558; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 197 AATAAACACCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGACATGGATGGATAA 256
Db 1 AATAAACACCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGACATGGATGGATAA 60
QY 257 TGAATAATTTCAACGTACCGTTACGTTGTACATGGCACCCTTCAGAAAGTACTCTGAG 316
Db 51 TGAATAATTTCAACGTACCGTTACGTTGTACATGGTACCGTTTCAGAAAGTACTCTGAG 120
QY 317 AGTTTCAGAGGCTGCAATACAAATATCTAGGTTGTCTGTTTCTGAAATATCTGAAATTTGCA 376
Db 121 AGTTTCAGAGGCTGCAATACAAATATCTAGGTTGTCTGTTTCTGAAATATCTGAAATTTGCA 180
QY 377 CACATGCTATGATTTACTCCAGATTCCTGGCAAGGTTATTGCCCTAGACACTTAATTACC 436
Db 181 TACATGCTATGATTTACTCCAGATTCCTGGCAAGGTTATTGCCCTAGACACTTAATTACC 240
QY 437 TGTGAAGCAATCATTCATATTTCTCCATGAACAGAGGGAATTCCTGTAGCTCCTCTCTGGGA 496
Db 241 TGTGAAGCAATCATTCATATTTCTCCATGAACAGAGGGAATTCCTGTAGCTCCTCTCTGGGA 300
QY 497 CTCATTCAGAGGTCATTTGTTGGTCTCCTTAGCCGATTTGGATTTTCATATTTGGG 556
Db 301 CTCATTCAGAGGTCATTTGTTGGTCTCCTTAGCCGATTTGGATTTTCATATTTGGG 360
QY 557 GGAGCTAGAAACTCATGGCTCGAACTTGACAGAGCAGCTTGAAACACACACTATATC 616
Db 361 GGAGCTAGAAACTCATGGCTCGAACTTGACAGAGCAGCTTGAAACACACTATATC 420
QY 617 TGCATGGAAGAGGCTTAAGCGGCAAACTAATGGAAGAAATGATAGTCACTGGCGACCGCA 676
Db 421 TGCATGGAAGAGGCTTAAGCGGCAAACTAATGGAAGAAATGATAGTCACTGGCGACCGCA 480
QY 677 ACAGCATCTAGTGCATGGCACCCTTATGAGTCCCTTAGGGACATTCGAGTAAAGCTTTT 736
Db 481 ACAGCATCTAGTGCATGGCACCCTTATGAGTCCCTTAGGGACATTCGAGTAAAGCTTTT 540
QY 737 GCATAATGGCATTTCCTACA 755
Db 541 GCATAATGGCATTTCCTACA 559

Search completed: July 9, 2004, 08:14:41
Job time : 4377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: July 9, 2004, 04:35:33 ; Search time 777 Seconds
(without alignments)
9885.766 Million cell updates/sec

Title: US-09-857-525C-1

Perfect score: 1576

Sequence: 1 gcacagagggaacatttgcg.....tcaaaaaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10C_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1423.4	90.3	1474	13	US-10-425-114-3817
2	1228	77.9	1844	13	US-10-425-114-1948
3	1226.4	77.8	2228	13	US-10-425-114-27250
4	1009	64.0	1387	13	US-10-425-114-33827
5	909	57.0	2980	17	US-10-437-963-2873
6	845.2	53.6	1200	13	US-10-425-114-6702
7	565.2	35.9	3077	13	US-10-424-599-44858
8	507	32.2	1507	13	US-10-425-114-29527
9	504.4	32.0	2315	17	US-10-424-599-79689
10	424.4	26.9	6747	17	US-10-437-963-93730
11	376	23.9	1112	17	US-10-437-963-22833
12	321.8	20.4	1225	13	US-10-424-599-101
13	310	19.7	721	13	US-10-425-114-30669
14	266.2	16.9	282	9	US-09-294-093B-135

15	241.6	15.3	1402	17	US-10-437-963-66100	Sequence 66100, A
16	219	13.9	1879	13	US-10-424-599-105	Sequence 105, App
17	217.6	13.8	1127	13	US-10-424-599-100	Sequence 100, App
18	188	11.9	1533	13	US-10-425-114-8857	Sequence 8857, App
19	176	11.2	1001	13	US-10-425-114-10694	Sequence 10694, A
20	175.2	11.1	365	12	US-09-732-627A-260	Sequence 260, App
21	163.2	10.4	954	13	US-10-425-114-2008	Sequence 2008, App
22	139	8.8	178	9	US-09-294-093B-4778	Sequence 4778, App
23	129.2	8.2	1140	13	US-10-424-599-44867	Sequence 44867, A
24	128.2	8.1	2362	13	US-10-425-114-31135	Sequence 31135, A
25	111.2	7.1	361	13	US-10-424-599-5459	Sequence 5459, App
26	100.2	6.4	411	13	US-10-424-599-44857	Sequence 44857, A
27	63.6	4.0	175	13	US-10-424-599-67266	Sequence 67266, A
28	58.8	3.7	373	13	US-10-424-599-59071	Sequence 59071, A
29	56.2	3.6	561	13	US-10-424-599-94913	Sequence 94913, A
30	54.4	3.5	917	13	US-10-424-599-44865	Sequence 44865, A
31	51.6	3.3	538	13	US-10-424-599-82474	Sequence 82474, A
32	50.6	3.2	11527	14	US-10-108-605-70	Sequence 70, Appl
33	47.2	3.0	517	17	US-10-021-323-11054	Sequence 11054, A
34	46.8	3.0	472	10	US-09-918-995-23770	Sequence 23770, A
35	45.6	2.9	485	17	US-10-021-323-6802	Sequence 6802, App
36	44.6	2.8	415	17	US-10-021-323-41	Sequence 41, Appl
37	44	2.8	200	17	US-10-021-323-16749	Sequence 16749, A
38	42.6	2.7	324	16	US-10-125-968-1261	Sequence 1261, App
39	42.4	2.7	1578	16	US-10-305-720-1113	Sequence 1113, App
40	42.4	2.7	1691	9	US-09-925-297-2	Sequence 2, Appl
41	42.2	2.7	391	13	US-10-424-599-59659	Sequence 59659, A
42	42.2	2.7	512	10	US-09-814-353-5072	Sequence 5072, App
43	42.2	2.7	512	10	US-09-814-353-11364	Sequence 11364, A
44	42	2.7	499	17	US-10-021-323-10189	Sequence 10189, A
45	42	2.7	1435	10	US-09-769-970-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-10-425-114-3817
; Sequence 3817, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3817
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700342215_FLI
US-10-425-114-3817

Query Match 90.3%; Score 1423.4; DB 13; Length 1474;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1457; Conservative 0; Mismatches 6; Indels 10; Gaps 2;

QY 86 CGAGTACAAATCTTTGTGGACGGGAATGGCGGATGATGCGGTCAACCTACCAATAC 145

Db 1 CGAGTACAAATCTTTGTGGACGGGAATGGCGGATGATGCGGTCAACCTACCAATAC 60

QY 146 TGGGGAGTTTGGCATAGTTAAACACTTTTACTTGAAGGGATATAACCAATAAACAC 205

Db 61 TGGGGAGTTTGGCATAGTTAAACACTTTTACTTGAAGGGATATAACCAATAAACAC 120

206 CTTATCAAGTCAAGCACACCTCGAGCAGGATGACATGATGTGGATTAATGAATTTT 265
 121 CTTATCAAGTCAAGCACACCTCGAGCAGGATGACATGATGTGGATTAATGAATTTT 180
 266 TCAACCTTACGCTTACGCTTGTGACATGSCACCGTTTCAGAGGTACTCTGAGAGTTTCAGA 325
 181 TCAACCTTACGCTTACGCTTGTGACATGSCACCGTTTCAGAGGTACTCTGAGAGTTTCAGA 240
 326 GGGTCAATACAAATATCTAGGTGTGCTGTTTCTGAAATATCTGAAATTTGCATACATCTA 385
 241 GGGTCAATACAAATATCTAGGTGTGCTGTTTCTGAAATATCTGAAATTTGCATACATCTA 300
 386 TGAATTTACTCCAGATCTCGAGAGGTATGSCCTAGACATTAATTTACTCTGAGACA 445
 301 TGAATTTACTCCAGATCTCGAGAGGTATGSCCTAGACATTAATTTACTCTGAGACA 360
 446 ATCAATTCATATCTCCATGAACAGGGGATTCCTGTAGCTCCCTCTGAGACTCAATTCAG 505
 361 ATCAATTCATATCTCCATGAACAGGGGATTCCTGTAGCTCCCTCTGAGACTCAATTCAG 420
 506 AGGTCAATTTGGTCTCTCTAGCCCAATTTGATCTCATATCTCATATTTGGGGAGCTAGA 565
 421 AGGTCAATTTGGTCTCTCTAGCCCAATTTGATCTCATATCTCATATTTGGGGAGCTAGA 480
 566 AACTCATGCTCGAATCTTGACAGAGAGAGCTTGAAACACACATATATCTCATGGA 625
 481 AACTCATGCTCGAATCTTGACAGAGAGAGCTTGAAACACACATATATCTCATGGA 540
 626 AGAGGCTAAGCGGCAAACTAATGGAAGAAATGATAGTCACTGSCGCGCAACAGCATCT 685
 541 AGAGGCTAAGCGGCAAACTAATGGAAGAAATGATAGTCACTGSCGCGCAACAGCATCT 600
 686 AGTGAATGCCACCCCTATGAGTCTCTGAGGACATTTGAGTAAGCTTTTGGAAATGG 745
 601 AGTGAATGCCACCCCTATGAGTCTCTGAGGACATTTGAGTAAGCTTTTGGAAATGG 660
 746 CATTTCTACAGTCCAGTATTTATTTATCATCATCATAGATGAGTCAATTCGCGAGTTATT 805
 561 CATTTCTACAGTCCAGTATTTATTTATCATCATCATAGATGAGTCAATTCGCGAGTTATT 720
 806 GCATCTTGCAATCTTTCTGGAATTTTGAATGATTTTGTAGATACCTTCAAAAACCTCAAC 865
 721 GCATCTTGCAATCTTTCTGGAATTTTGAATGATTTTGTAGATACCTTCAAAAACCTCAAC 780
 866 TGGTAATTTGCCATTTCTGGAACCAACAGTGTGCTCCATTCGCTGGTTCCTGGGTTC 925
 781 TGGTAATTTGCCATTTCTGGAACCAACAGTGTGCTCCATTCGCTGGTTCCTGGGTTC 840
 926 GAAATTTGGTGTATCTGAACAGTGTGCTCCATTTGGTGTATGTTGCGACCTTAATGCTTACCTAG 985
 841 GAAATTTGGTGTATCTGAACAGTGTGCTCCATTTGGTGTATGTTGCGACCTTAATGCTTACCTAG 900
 986 CTCTGCGCTTAAATGTTGTTTCAAGCTGAGTAAGCTCAATACCAATTTGGATGACAA 1045
 901 CTCTGCGCTTAAATGTTGTTTCAAGCTGAGTAAGCTCAATACCAATTTGGATGACAA 960
 1046 CGACTCCCTGTTGACATTTACTCTAGAAGTGAATCAGACCTTACGAAAGACAGAGT 1105
 961 CGACTCCCTGTTGACATTTACTCTAGAAGTGAATCAGACCTTACGAAAGACAGAGT 1020
 1106 CTACACATGTTTGGGTGATGAGATGACATTTACACAGGCTTTGAGCTTGGACAAGA 1165
 1021 CTACACATGTTTGGGTGATGAGATGACATTTACACAGGCTTTGAGCTTGGACAAGA 1080
 1166 TGGCAATACACCTTTTGGATTTCTTAAAGCGGAGAGATGCGAGATGTCCTCGGTCTGA 1225
 1081 TGGCAATACACCTTTTGGATTTCTTAAAGCGGAGAGATGCGAGATGTCCTCGGTCTGA 1140
 1226 TCCCTTTGCTGAAGGTGATGAGGAGTCTGGCTTAATCTCTGGGTGCGGCGGTGTTCTATGT 1285
 1141 TCCCTTTGCTGAAGGTGATGAGGAGTCTGGCTTAATCTCTGGGTGCGGCGGTGTTCTATGT 1200
 1286 GGAAGCTGGGAGCAACCTGTGGAGGCGCATCATATCACTAAGTGAATTTTCAAGTTCTT 1345

1201 GGAAGCTGGGAGCAACCTGTGTGAGGGCATATATCACTAAGTACATTTTCAAGTTCTT 1260
 1346 GCTGAGCTTCTGAGGAGAGAGGAGCATGCGGAGTGTGTTTTTAACTAGTAGCAACCTCTG 1405
 1261 GCTGAGCTTCTGAGGAGAGAGGAGCATGCGGAGTGTGTTTTTAACTAGTAGCAACCTCTG 1320
 1406 GGCCTTTGGTCTCTGCGGATAATTTGGCTATGTCAGGACTCAGGGTTCATGCTGTGA 1465
 1321 GGCCTTTGGTCTCTGCGGATAATTTGGCTATGTCAGGACTCAGGGTTCATGCTGTGA 1372
 1466 AAGTTCCTTTTGTGCTTGTGTTTTTACATTTGGCCCTCCCGTTGTTGTTAAACAAACAT 1525
 1373 AAGTTCCTTTTGTGCTTGTGTTTTTACATTTGGCCCTCCCGTTGTTGTTAAACAAACAT 1430
 1526 ACATCAGGTGATTTTATCTATTTGGATTGTTTC 1558
 1431 ACATCAGGTGATTTTATTTATTGATTGTTTC 1463

RESULT 2
 US-10-425-114-1948
 ; Sequence 1948, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 1948
 ; LENGTH: 1844
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700172325_FLI
 US-10-425-114-1948

Query Match 77.9%; Score 1228; DB 13; Length 1844;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1279; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

8 GGAACATTTGCCGATGCTCTCTATCGAAGGCTGCCACCTGATTTTCAGGCTATTTGCAG 67
 205 GGAACATTTGCCGATGCTCTCTATCGAAGGCTGCCACCTGATTTTCAGGCTATTTGCAG 264
 68 TCTGCTCTCTCGGATTCACGAGTACAAATTTCTTTGTGACGGGGAATGGCGGATGATGA 127
 265 CTTGCTCTCGGATTCACGAGTACAAATTTCTTTGTGACGGGGAATGGCGGATGATGA 324
 128 GCTGCACTTACCATATCTGGGAGATTTGCGATAGTTAAACACTTTACTTTCAGAGGA 187
 325 GCGCCAACTTACTATATCTGGGAGTTCGGTATAGTTAAACACTTACTTTCAGAGGA 384
 188 ATATAACCAATTAACACCTTATCAAGTCCAAAGCACACCTGGAAGAGGATGAACATGGA 247
 385 ATTATCAATTAATTAATGCTTATTAATCAAGCACACCTGGAAGAGGATGAACATGGA 444
 248 TGTGATAATGAATTTTCAACGTAACGTTACGTTGTGTCAGATGGCAACGTTTCAGAGG 307
 445 TGTGATAATGAATTTTCAACATACGTTTACATTTGTCGATGTTACCATTTCCAGAGG 504
 308 TACTCTGAGGTTTCAGAGGCTGCAATACAAATATCTAGGTGCTGTTCTGATATCT 367
 505 TACTGTGAGGTTTCAGAGGCTGCAATACAAATATCTAGGTGCTGTTCTGATATCT 564

368 GAATTGACATCATGCTATGATTTACTCCAGATTCTGGCAAGGTTATTSCCTAGACAT 427
Db
565 CAATTGACATCATGCTATGATTTACTCCAGATTCTGGCAAGGTTATTSCCTAGACAT 624
Qy 428 TAATTACCTGTGAAGCAATCAATTCATATCTCCATGAAAGGAGGATTCCTAGCTCC 487
Db 625 TAATTACCTGTGAAGCAATCTTTTCATATCTCCATGAAAGGAGGATTCCTAGCTCC 684
Qy 488 TCTGTGGGACTCATTCAGAGGCTCAATTTGTGGCTCTCTAGCCCATTTGGATTTCATCT 547
Db 685 TCTGTGGGACTCATTCAGAGGCTCAATTTGTGGCTCTCTAGCCCATTTGGATTTCATCT 744
Qy 548 CATATTGCGGAGCTAGAACTCATGGCTCGAATCTTGCAGAGAGAGAGAGCTTGAACACA 607
Db 745 TATATTGCGGAGCTAGAACTCATGGCTCGAATCTTGCAGAGAGAGAGAGCTTGAACACA 804
Qy 608 CACTATATCTGATGAAAGAGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCT 667
Db 805 CACTATATCTGATGAAAGAGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCT 864
Qy 668 GCGACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGAGCATTTGAGT 727
Db 865 GCGACGACATCAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGAGCATTTGAGT 924
Qy 728 AAAGCTTTTGCAGAAATGGCAATTTCTACAGTCCAGTTCATTTATTCATCATCATCATGG 787
Db 925 AAAGCTTTTGCAGAAATGGCAATTTCTACAGTCCAGTTCATTTATTCATCATCATCATGG 984
Qy 788 ATCAATTCGCGAGCTTATGATCTTGCATCATCTTCTGAAATTTGAAATGATTTGAGT 847
Db 985 ATCAATTCGCGAGCTTATGATCTTGCATCATCTTCTGAAATTTGAAATGATTTGAGT 1044
Qy 848 ATACTTCAAAATCTCAACTGTAATTTGCTATTTCTGAAACCAACAGTGTGCTCCATTC 907
Db 1045 GTATTTTAAATCTCAACTGTAATTTGCTATTTCTGAAACCAACAGTGTGCTCCATTC 1104
Qy 908 GCTGGGTTCTCGGTTCCGAAATTTGGTGTATCTGAAAGTGTGCTCAATTTGTTGG 967
Db 1105 GCTGGGTTCTCGGTTCCGAAATTTGGTGTATCTGAAAGTGTGCTCAATTTGTTGG 1164
Qy 968 ACCTAATGCTCATTTAGCTCTGCTTCAATGATGATGATGATGATGATGATGATGAT 1027
Db 1165 ACCTAATGCTCATTTAGCTCTGCTTCAATGATGATGATGATGATGATGATGATGAT 1224
Qy 1028 ACCAATGCTGATGAAAGCACTCCCTGCTGACATCTACTAGAAAGTGCATCATCAGC 1087
Db 1225 ACCAATGCTGATGAAAGCACTCCCTGCTGACATCTACTAGAAAGTGCATCATCAGC 1284
Qy 1088 CCTAGCAAAAGCAAGGCTTACACATGTTGGCTGATGATGATGATGATGATGATGATG 1147
Db 1285 CCTAGCAAAAGCAAGGCTTACACATGTTGGCTGATGATGATGATGATGATGATGATG 1344
Qy 1148 TTTGAGCTTGGACAGATGCCAATCACTTTTGGATTTCTTAAAGGCGGAGATGCCA 1207
Db 1345 TTTGAGCTTGGACAGATGCCAATCACTTTTGGATTTCTTAAAGGCGGAGATGCCA 1404
Qy 1208 GATGTGCTCCGCTCTGATCCTTTGCTGAAGGATGATGAGGAGTCTGCTAATCTCGGGT 1267
Db 1405 GATGTGCTCCGCTCTGATCCTTTGCTGAAGGATGATGAGGAGTCTGCTAATCTCGGGT 1464
Qy 1268 GCGGCGGCTTCAATGTGAAAGCTGGAGCAACGCTGTGGAGGAGTATCATCATTAAG 1327
Db 1465 GCGGCGGCTTCAATGTGAAAGCTGGAGCAACGCTGTGGAGGAGTATCATCATTAAG 1524
Qy 1328 TGACATTTCAAGTTCTTCTGAGCTTGTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAG 1371
Db 1525 TGACATTTCAAGTTCTTCTGAGCTTGTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAG 1568

RESULT 3

US-10-425-114-27250

; Sequence 27250, Application US/10425114

; Publication No. US20040034889A1

GENERAL INFORMATION:

; APPLICANT: Liu, Jindong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 27250

; LENGTH: 2228

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: L1B4729-018-A12_FLI

US-10-425-114-27250

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Query Match

77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```

QY 728 AAGCTTTTGCAGAAATGGCAATTTCTACAGTCCAGTATTTATTCATCATCATCAGATGG 787
DB 1308 AAGCTTTTGCAGAAATGGCAATTTCTACAGTCCAGTATTTATTCATCATCATCAGATGG 1367
QY 788 ATCAATCCCGAGTATGATCTGTCATCACTTTCTGGAATTTGAAATGTTATTTCTAG 847
DB 1368 ATCAATCCCGAGTATGATCTGTCATCACTTTCTGGAATTTGAAATGTTATTTCTAG 1427
QY 848 ATACTTCAAAATCAAACTGGTAAATTTGGCTATTTCTGAAACCAACAGCTGGCTCCATCC 907
DB 1428 GTATTTTAAAACTCAATGGTAAATTTGGCTATTTCTGAAACCAACAGCTGGCTCCATCC 1487
QY 908 GCTGGGTTCTGGGTTCCGAAATTTGGTGAATCTGAAACAGTGGTCCATGGCTATGTTGG 967
DB 1488 GCTGGGTTCTGGGTTCCGAAATTTGGTGAATCTGAAACAGTGGTCCATGGCTATGTTGG 1547
QY 968 ACCTAATGCTCACTAGCTTGGCTTACATGTTGCTTCAAGCTGGATGAAGCTCAAT 1027
DB 1548 ACCTAATGCTCACTAGCTTGGCTTACATGTTGCTTCAAGCTGGATGAAGCTCAAT 1607
QY 1028 ACCAATTTGGATGACCAACAGCTCCCTGCTTGACACTTACTCTAGAAAGTGACATCAGC 1087
DB 1608 ACCAATTTGGATGACCAACAGCTCCCTGCTTGACACTTACTCTAGAAAGTGACATCAGC 1667
QY 1088 CCTAGCAAAAGAGTCTACACATGTTTGGCTGGATGAGATGACATCAGCAGC 1147
DB 1668 CCTAGCAAAAGAGTCTACACATGTTTGGCTGGATGAGATGACATCAGCAGC 1727
QY 1148 TTTGACGCTTGGCAAGATGCCATACACCTTTTGGATTTCTTAAAGCCAGAGATGCCA 1207
DB 1728 TTTGACGCTTGGCAAGATGCCATACACCTTTTGGATTTCTTAAAGCCAGAGATGCCA 1787
QY 1208 GATGCTGCTCCGGTCTGATCCTTTGCTGAAGTGTATGAGAGCAGCTGGCTAATCTCGGGT 1267
DB 1788 GATGCTGCTCCGGTCTGATCCTTTGCTGAAGTGTATGAGAGCAGCTGGCTAATCTCGGGT 1847
QY 1268 GCGCGGGTGTCTATTTGGAGCTGGAGCGAAGCTGTGAGGCGCATCATCACTTAG 1327
DB 1848 GCGCGGGTGTCTATTTGGAGCTGGAGCGAAGCTGTGAGGCGCATCATCACTTAG 1907
QY 1328 TGCAATTTCAAGTTCTTGGTGAAGCTGTGAGGAGAGAGGAGCA 1371
DB 1908 TGCAATTTCAAGTTCTTGGTGAAGCTGTGAGGAGAGAGGAGCA 1951

```

RESULT 4

```

US-10-425-114-33827
; Sequence 33827, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33827
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMPLM017174G09_FLI
US-10-425-114-33827

```

Query Match 64.0%; Score 1009; DB 13; Length 1387;
 Best Local Similarity 94.2%; Pred. No. 8.5e-296;
 Matches 1048; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```

QY 259 AAAATTTTCAACGTACCGTTTACGGTTTCAAGTGCACCGTTTCAAGAGTACTCTGAGAG 318
DB 1 AAAATTTTCAACGTACCGTTTACGGTTTCAAGTGCACCGTTTCAAGAGTACTCTGAGAG 60
QY 319 TTTTCAGAGGCTGCAATACAAATATCTAGGTTCTCGTTTCTGAAATATCTGAATTTGCAATA 378
DB 61 TTTTCAGAGGCTGCAATACAAATATCTAGGTTCTCGTTTCTGAAATATCTGAATTTGCAATA 120
QY 379 CATGCTATGATTTTACTCCAGATTTCTGGCAAGTTTATTTGCCCTAGACATTAATTTTACCTG 438
DB 121 CATGCTATGATTTTACTCCGATTTCTGGCAAGTTTATTTGCTCTCGACATTAATTTTACCTG 180
QY 439 TGAAGCAATCATTTCCATATTTCTCATGAAACAGGGAATCTCTGAGTCTCTCTCTGGGACT 498
DB 181 TGAAGCAATCTTTTTCATATTTCTGCATGAACAGGGAATCTCTGAGTCTCTCTCTGGGACT 240
QY 499 CATTCAGAGGCTCAATTTGTTGGTCTCTTACCCATTTGGATTTTCATATCTCATATTTGCGGG 558
DB 241 CATTCAGAGGCTCAATTTGTTGGTCTCTTACCCATTTGGATTTTCATATTTTATTTGCGGG 300
QY 559 AGCTTAGAATCTCATGCTCGAATTTTGAACAGAGGAGCTTTGAAACACACACTATATCTG 618
DB 301 AGCTTAGAATCTCATGCTCGAATTTTGAACAGAGATCAGTTTGAACACACACTATATCTG 360
QY 619 CATGGAAGAGGCTTAAGCGGCAAACTAATGGAAGAAATGATGATGAGTGGCGAGCCGCAAC 678
DB 361 CATGGAAGAGGCTTAAGCGGCAAACTTGTGGAAGAAATGATGATGAGTGGCGAGCCGCAATC 420
QY 679 AGCATCTAGTGCATGCGCAACCCCTTATGAGTCTCTTGGAGGACATTTGAGTAAAGCTTTTGC 738
DB 421 AGCATCTAGTGCATGCGCAACCCCTTATGAGTCTCTTGGAGGACATTTGAGTAAAGCTTTTGC 480
QY 739 AAAATGGCAATTTTCAAGTGCAGTATTTATTCATCATCATCATCATCATCATCATCATCAT 798
DB 481 AAAATGCAATTTTCTACAGTGCAGTATTTATTCATCATCATCATCATCATCATCATCATCAT 540
QY 799 AGTTATTTGCACTTTTGGCATCACTTTCTGGAATTTTGAATGATTTTGTAGATTAATTTCAAAA 858
DB 541 AGTTATTTGCACTTTTGGCATCACTTTCTGGAATTTTGAATGATTTTGTAGTATTTTAAAA 600
QY 859 ACTCAACTGGTAAATTTTGGCTTATTTCTGAACCAACAGTGTCTCCATTTCCGCTGGGTTCT 918
DB 601 ACTCAACTGGTAAATTTTGGCTTATTTCTGAACCAACAGTGTCTCCATTTCCGCTGGGTTCT 660
QY 919 GGGTTCCGAAAATTTGGTGCATCTGAAACAGTGTCTCATTTGGCTATTTGTGGAGCTAATGCT 978
DB 661 GGGTTCCGAAAATTTGGTGCATCTGAAACAGTGTCTCATTTGGCTATTTGTGGAGCTAATGCT 720
QY 979 CACTTAGCTCTGCTTAAACATTTTGGTTTCAAGCTGGAGTAAAGCTCAATACCAATTTTGG 1038
DB 721 CACTTAGCTCTGCTTAAACATTTTGGTTTCAAGCTGGAGTAAAGCTCAATACCAATTTTGG 780
QY 1039 ATGACAAAGCTCTCTGCTTGCACATTTACTCTAGAAAGTGACATCAAGCTTACGAAAG 1098
DB 781 ATGAAAACGACTCTCTGCTTGCACATTTACTCTAGAAAGTGACATCAAGCTTACGAAAG 840
QY 1099 ACAAGGCTCTACACATTTTGGCTGGATGAGATGACCACTTACAGGCTTTGAGCTTG 1158
DB 841 ACAAGGCTCTACACATTTTGGCTGGATGAGATGACCACTTACAGGCTTTGAGCTTG 900
QY 1159 GACAAGATGCCAATACACCTTTTGGATTTCTTAAAGCGCCAGAGATGCCAGATGTCCTCC 1218
DB 901 GACAAGATGCCAATACACCTTTTGGATTTCTTAAAGCGCTCAGAGATGCCAGATGTCCTCC 960
QY 1219 GGTCTGATCTCTTGTCTGAAGTGTATGAGAGCACTGGCTTAATCTCTGGGCTGGCGGGTGT 1278
DB 961 GGTCTGATCTCTTGTCTGAAGTGTATGAGAGCACTGGCTTAATCTCTGGGCTGGCGGGTGT 1020
QY 1279 TCATTTGTGGAAGCTGGGAGGAAACGTTGTGAGGAGGATCATATCACTAAGTGAATTTTCA 1338
DB 1021 TCATAGTAGAAGCTGGGAGGAAACGTTGTGAGGAGGATTTATATCACTAGTGTATTTTCA 1080

```

QY 1339 ACTTCTGCTGAGCTTGTGAGGAGAGAGCA 1371
Db |||||
1081 AGTTCTGCTGAGCTTGTGATGAGAGAGGA 1113
|||
RESULT 5
US-10-437-963-2873
; Sequence 2873, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 2873
; LENGTH: 2980
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT453c_102608C.1
US-10-437-963-2873
Query Match 57.7%; Score 903; DB 17; Length 2980;
Best Local Similarity 82.6%; Pred. No. 4.1e-265;
Matches 1041; Conservative 0; Mismatches 220; Indels 0; Gaps 0;
QY 5 GAGGGAACATTTGCGGATGCTCTCTATCGAAGGCTGCCCACTGTATTTTCNGGCTATTTG 64
Db |||||
681 GACAGACATTTGCCAATGCTCTCTGTTGAAGTTGTCCACCGTTTTCNGGCCATATG 740
|||
QY 65 GAGTCTGCTCTGGGATTCAGAGTACAAATCTTTTGGACGGGATCGGGCATGA 124
Db |||||
741 CAGTTTGTCCCTGGGATTCATCAGTACAAATTTTGTGTGATGGGATGGCGCACGA 800
|||
QY 125 TCAGGCTCAACTACCATATCTGCGGAGTTTGGCATAGTTTAAACACACTTTTACTTTGACAAG 184
Db |||||
801 TGAGGGCAACTACCATACAGGAGCATAGTGTGTGTTTAACTTTTACTTTGACTAG 860
|||
QY 185 GGAATATAACCAATAAACAACCTTATCAAGTCCAGCACACCTGGGAAGCAGGATGAACAT 244
Db |||||
861 GGAATTTGACCAAGATAAATAACAATATTGAGCCCTAGTACACCTGGGAAGTAGGATGAACAT 920
|||
QY 245 GGAATGGATATGAAATTTTCAAGTACGTTAGTGTGATGAGGACCGTTTCAGA 304
Db |||||
921 GGAATGGACAAAGCAATTTTCAAGTACTGTTTCTTGTCTGATGGCATTTATTCAGGA 980
|||
QY 305 AGGTAATCTGAGAGTTTCAGAGGCTCAATACAAATATCTAGTGTGCTGTTTCTCAATA 364
Db |||||
981 AGGTTCTGAGAGTTTCAGAGGAGCTATACAGATCTCTAGTGTGCTGTTGAGCGATTT 1040
|||
QY 365 TCTGAATTTGCATACATGCTATGATTTACTTCCAGATCTTGGCAAGGTTATTTGCCCTAGA 424
Db |||||
1041 TCTGAATGACAAACTGGGTATGATTTACTTCCAGATCTTGGCAAGGTCATTTGCTCTAGA 1100
|||
QY 425 CATTAATTTACCTGTGAAGCAATCATCTATCTCCATGACAGGAGGATTCCTGTAGC 484
Db |||||
1101 CCAATTTGCTGTGAAGCAATCTTTTCAATTTCTTATGATGATGATGATGATGATGATGAT 1160
|||
QY 485 TCTCTCTGGGACTCATTCAGAGGTCATTTTGTGTCTCTCTTCTAGCCCATTTGATTTCTAT 544
Db |||||
1161 ACCTCTATGGGATTCATTCAGGGGCCAGTTTGTGGCCCTTTTGTAGCCCATCTGATTTTAT 1220
|||
QY 545 ACTCATATTCGGGAGCTAGAAATCATGCTCGAATTTGACAGAGAGAGCAGCTTGAAC 604
|||

Db |||||
1221 ACTCATATTTGAGAGAGCTGGAACCTCATGCTCTCAATCTCACAGAGAGAGCTTGAAC 1280
QY |||||
605 ACACACTATATCTGCATGGAAGAGGCTAAGCGGCAAACTAAATGGAGAAATGATATCA 664
Db |||||
1281 ACATACATATCTGCTGGAAGAGGCGCAGGCTCAAACTTATGCGCAAAACGAGGCTTC 1340
|||
QY 665 GTGGCGACCGCAACAGCATCTAGTGCAATGCAACCCCTTATGAGTCTCTGAGGAGCATTCG 724
Db |||||
1341 TTGGAGGCAAAATCACCATTAGTTTATGCGCACCCCTTATGATCCCTGAGGAAATTCG 1400
|||
QY 725 AGTAAAGCTTTTGCAAAATGGCATTTCTACAGTGGCGAGTTATTTATCATCATCAGA 784
Db |||||
1401 TATGAAGATACCTGCAAAACGGTGTCTACCGTTCCAAATATGTTTCTCTCATCACCAGA 1460
|||
QY 785 TGGATCATTTCCCGCAGTTATGCAATCTGCAATCACTCTCTGGAATTTGGAATGATTTG 844
Db |||||
1461 TGCTCATATCCCAATTTGTCATCTGCTCCCTTCTGGAATTTGGAATGATTTG 1520
|||
QY 845 TAGATACCTTCAAAATCTCAACTGGTAATTTGCTCTTCTGAAACCAACAGTGTCTCCAT 904
Db |||||
1521 TAGATATTTTAAATTTCTCAAGGTAATTTACCTATTTTGGCCAACTCTATGCCAAT 1580
|||
QY 905 TCCGCTGGGTTCTTGGGTTCCGAAATTTGGTGAATCTGAACAGCTCGTCCATTTGCTATGTT 964
Db |||||
1581 TCCTCTGGGTACCTGGGTTCCAAAATTTGGTGAATCTCTATGCGCTCCATTTGGCTATGTT 1640
|||
QY 965 GGGACCTAATGCTCTCACTTAGCTCTGCCCTTACAGCTTGGTTCAAGCTGGAGTAAGCTC 1024
Db |||||
1641 GCGGCTTACACATCTCTTAGCGCTGCCCTGGAATTTGCTGTTCAAGCTGGTGGAGCTC 1700
|||
QY 1025 AATACCAATTTGGATGACAAAGCTCTCTGTTGACACTTACTCTAGAAAGTGACATCAC 1084
Db |||||
1701 AATACCAATTTGGATGACAAAGCTCTCTGTTGACACTTACTCTAGAAAGTGACATCAC 1760
|||
QY 1085 AGCCTTAGCAAAAGAGAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCA 1144
Db |||||
1761 AGCTCTTGCAAAAGAGAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCA 1820
|||
QY 1145 GCTTTGCGAGCTTGGACAGATGCCAATTCACCTTTTGGATTTCTTAAACGCCACAGATG 1204
Db |||||
1821 GCTTCTGAGCTTGGACAGATGCCAATTCACCTTTTGGATTTCTTAAACGCCACAGATG 1880
|||
QY 1205 CCAGATGTGCTCGGCTCTGATCTTTTGGTGAAGGTGATGAGGAGTGGCTTAATCTGCG 1264
Db |||||
1881 TCAGATGTGCTCGGCTCTGATCTTTTGGTGAAGGTGATGAGGAGTGGCTTAATCTGCG 1940
|||
QY 1265 G 1265
Db 1941 G 1941

RESULT 6

US-10-425-114-6702
; Sequence 6702, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6702
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: 700575557_FLI
US-10-425-114-6702

Query Match 53.6%; Score 845.2; DB 13; Length 1200;
Best Local Similarity 94.8%; Pred. No. 5.9e-246;
Matches 874; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```

QY 450 TTCCATATTCCTCCATGAACAGGGGATTCCTGTAGCTCTCTCTGGGACTCATTGAGGT 509
DB 2 TTTTCATATTCCTCCATGAACAGGGGATTCCTGTAGCTCTCTCTGGGACTCATTGAGGT 61
QY 510 CAATTTGTTGGTCTCTTAGCCCATTTGGATTTCTATCTATATTCCTGGGAGCTAGAACT 569
DB 62 CAATTTGTTGGTCTCTTAGCCCATTTGGATTTCTATCTATATTCCTGGGAGCTAGAACT 121
QY 570 CATGGCTCGAATCTTGACAGAGAGCAGCTTGAAACACACACTATATCTGATGAAAGAG 629
DB 122 CATGGCTCGAATCTTGACAGAGAGCAGCTTGAAACACACACTATATCTGATGAAAGAG 181
QY 630 GCTAAGCGGCAAACTAATGAAGAAATGATGATGATGATGATGATGATGATGATGATG 689
DB 182 GCTAAGCGGCAAACTTGTGAAGAAATGATGATGATGATGATGATGATGATGATGATG 241
QY 690 CATGCCACCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
DB 242 CATGCCACCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
QY 750 TCTACAGTGCAGTATTTATTTATTCATCATCATCATCATCATCATCATCATCATCATCAT 809
DB 302 TCTACAGTGCAGTATTTATTTATTCATCATCATCATCATCATCATCATCATCATCATCAT 361
QY 810 CTTCGATCATCTTCTGAAATTTGAAATGATGATGATGATGATGATGATGATGATGATG 869
DB 362 CTTCGATCATCTTCTGAAATTTGAAATGATGATGATGATGATGATGATGATGATGATG 421
QY 870 AATTTGGCTATCTTGAAACCAAGTGTGCTCCATTCGGTGGGTCTCGGGTTCGGAA 929
DB 422 AATTTGGCTATCTTGAAACCAAGTGTGCTCCATTCGGTGGGTCTCGGGTTCGGAA 481
QY 930 AATTTGGTGAATCTGAACAGTGTGCTCCATTCGGTGGGTCTCGGGTTCGGAA 989
DB 482 AATTTGGTGAATCTGAACAGTGTGCTCCATTCGGTGGGTCTCGGGTTCGGAA 541
QY 990 GCGCTTAACATGTTGTTCAAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1049
DB 542 GCGCTTAACATGTTGTTCAAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 601
QY 1050 TCCCTGCTTGACACTTACTCTAGAGTGACATCACAGCCCTAGCAAAAGACAAGGTCTAC 1109
DB 602 TCCCTGCTTGACACTTACTCTAGAGTGACATCACAGCCCTAGCAAAAGACAAGGTCTAC 661
QY 1110 ACACATGTTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1169
DB 662 ACACATGTTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721
QY 1170 AATACACCTTTTGAATTTTAAAGGCGCAGAGATGCGAGATGCGCTCCGCTGATCCT 1229
DB 722 AATACACCTTTTGAATTTTAAAGGCGCAGAGATGCGAGATGCGCTCCGCTGATCCT 781
QY 1230 TTGCTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1289
DB 782 TTGCTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
QY 1290 GCTGGAGGCAAACTGTTGGAGGGGATGATGATGATGATGATGATGATGATGATGATG 1349
DB 842 GCTGGAGGCAAACTGTTGGAGGGGATGATGATGATGATGATGATGATGATGATGATG 901
QY 1350 AGCTTGTGAGGAGGAGGACA 1371
DB 902 AGCTTGTGATAGAGAGAGGA 923

```

RESULT 7

US-10-424-599-44858
Sequence 44858, Application US/10424599
Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 44858
LENGTH: 3077
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_140506C.1
US-10-424-599-44858

Query Match 35.9%; Score 565.2; DB 13; Length 3077;

Best Local Similarity 65.9%; Pred. No. 2.1e-160;
Matches 883; Conservative 0; Mismatches 448; Indels 9; Gaps 4;

```

QY 9 GAACATTTGCGGATGCTCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAGT 68
DB 1019 GAACATTTGCGGATGCTCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAGT 1078
QY 69 CTGCTCTCTGGGATTCACGAGTACAAATCTCTTCTGACCGGGGATGCGGCATGATGAG 128
DB 1079 TTGCTACCCGGTACCACATCAGTACAAATCTCTTCTGATGAGGATGCGGCATGATGAG 1138
QY 129 CGTCAACCTTACCATATCTGCGGAGTTTGGCATAGTTAAACACACTTACTTTCACAAGGGA 188
DB 1139 CATCAACCTTATGCTAGCTGGAGATATGGATAGTCTAACACTGT-CTTATTGGGCCAC 1197
QY 189 TATAACCAATTAACACCTTATCAAGTCCAAAGCACCTGGAAGCAGATGAAATGAT 248
DB 1198 TGATCTCTAATTAACATACCTGTTTACCTCCAG--ACGTTGCTTCTGGAATAGCATGAT 1255
QY 249 GTGATATGAATTTTCAACGTAACGTTTACGTTTGTGATGATGATGATGATGATGATG 308
DB 1256 GTGATATGATGCTTTTCCCGGATGCGCGGTTGACCGATGATGATGATGATGATGATG 1315
QY 309 ACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGTGTCTGTGTTCTGAAATATCT 368
DB 1316 CTGCAAGATATCAGATACCTGATGTACAAATATCCCGTCAGCGTATTTCTGCATTCTA 1375
QY 369 AATTTGATACATGCTATGATTTACTCCAGATTTCTGCAAGGTTATTCCTAGACATT 428
DB 1376 TCTTCACACCCGTTATGAAATTTACTTCCGAGTCAGGCAAGGTTGTGCTTTGGATGTT 1435
QY 429 AATTTACTCTGGAAGCAATCAATTCCTATTTCTCATGAAACAGGGGATTTCTGTAGCTCT 488
DB 1436 GATCTACCACTGAAACAGGCAATTCATATATTCATGAGCAGGAGGTTTTCATGGCTCT 1495
QY 489 CTCTGGAGCTCATTCAGAGGTCATTTGTTGGTCTCTCTAGCCCATTTGGATTTTCATCTC 548
DB 1496 CTTTGGAGCTTCTGCAAGGGCAATTTGTTGGTCTCTCTAGTGTCTTGGATTTTATTTTA 1555
QY 549 ATATTTGGGAGCTGAGAACTCATGCTCGCACTGCAAGCAGCTGCAAGAGCAGCTTGAACAC 608
DB 1556 ATTTTAAGAGAGCTGGGGAATCATGGATCCAAATCTGACAGAGAGGAGCTTGAACACAT 1615
QY 609 ACTATATCTCATGGAAGAGGCTTAGCGGCAAACTAATGGAAGAAATGATGATGATGATG 667
DB 1616 ACCATATCAGCTTGGAGAGGAGGAAATCGTATCTAAATAGACAGAACTATGGACATGA 1675
QY 668 --GCGACCGCAACAGCATCTAGTGTGATGCGCCCTTATGATGCTTTGAGGGGACATTGCA 725

```

Db 1676 ACTGCATTTTCAAGATGTTTATTCACGAGCGGCCATATGATATCTGAAAGATATTGGC 1735
QY 726 GTAAAGCTTTTGCATAATGCAATCTACAGTGCAGGTATTTATTTATTCATCATCATCAGAT 785
Db 1736 ATGAGATCTTGCMAAGAGGAGTTTCAACAGTTCCTTATATCCATTCATCTTCGAAGAT 1795
QY 786 GGAATCATCCCGAGTTATTCATCTTGCATCACTTCTCTGGAATTTTGAATGATTTGT 845
Db 1796 GCTTCATTTCCACAGTTACTACATCTGCTTCACTTTCAGGAATCTTAAATGCAATGT 1855
QY 846 AGATATCTTCAAAACTCAAGTGTAAATTCGCTATTTCTGAACCAACAGTGTCTCAAT 905
Db 1856 AGGTATTTTAGGCACGCTCTAGTCTCTTGCCTGACTTCAACTTCCAACTCTGCAATA 1915
QY 906 CCCTGGGTTTCTGGGTTCCGAAATTTGGTGATCTGAACAGTGTCTCAATGGCTATCTTG 965
Db 1916 CTTGTGGCAGGTGGTGGCCCAAAATTTGGGGAATCAATCGCGGCCCTTAGCAATGTTG 1975
QY 966 CGACCTAATGCCTACCTTAGCTCTGCGCTTAAACATGTTGGTTCAAGCTGGAGTAAGCTCA 1025
Db 1976 AGACCAACCGCTTCTCTTGGCTCAGCGCTTAAATTTATTTAGTTCAAGCCCAAGTAAGCTCA 2035
QY 1026 ATACCAATTTGATGATGACAGACTCCCTGCTTGACACTTACTCTAGAAGTGACATCACA 1085
Db 2036 ATACCAATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2095
QY 1086 GCCTTAGCAAAAGACAAAGTCTTACACACATGTTTGGGCTGATGATGATGATGATGATGATGAT 1145
Db 2096 GCTTTGGCAAGAACAGACAGACATACACATATTAATTTGACGAAATGACTGTTTCATCAG 2155
QY 1146 GCTTTGCAGCTTGGACAGATGCCATACACCTTTTGGATTTCTTAAGCGCCAGAGATGC 1205
Db 2156 GCTTTGCAGTTGGCCAGGACGATATAGTCCCTATG---AGCTTAGAAGTCAAGATGT 2212
QY 1206 CAGATGTGCTCGGTTCTGATCTTGTCTGAAGTGTGAGGAGCTGGCTATCTCTGGG 1265
Db 2213 CAGATGTGTTGGTTCTGATCTCTGATPAAAGTGAAGGAGCTTGGCAATCCAGGT 2272
QY 1266 GTGGGCGGCTGTTCAATTTGGAAGCTGGGAGCAAAAGTGTGGAGGCGCATCATCACTA 1325
Db 2273 GTGAGCGGCTGTGATCTGTTGAAGCTGGCAGCAAGCGTGTAGAGGCAATGTTTCATTG 2332
QY 1326 AGTGACATTTTCAAGTCTT 1345
Db 2333 AGTGACATTTCAAGTCTT 2352

RESULT 8

US-10-425-114-29527
; Sequence 29527, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 28
; CURRENT FILING DATE: 2003-04-28
; SEQ ID NO 29527
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURES:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY057G02_FLI
US-10-425-114-29527

Query Match

32.2%; Score 507; DB 13; Length 1507;

Best Local Similarity 67.6%; Pred. No. 6.8e-143;
Matches 743; Conservative 0; Mismatches 350; Indels 6; Gaps 2;
QY 250 TGGATATGAAATTTTCAAGTACGGTTAGTGTGATGAGTGGACCGTTTTCAGAGGTA 309
Db 1 TGGATATGATGCTTTTGGCGGAATGCGCCGGTTGACCGATGGTACTTTGAGTGAAGTGC 60
QY 310 CTCTGAGAGTTTTCAGAGGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTGA 369
Db 61 TGCCAAGATATCAGATACATGATGATCAAAATATCCCGTCAGCGTATTTCTGCAATTTCTAT 120
QY 370 ATTTGCAATGATGATGATTTACTCCAGATTTCTGCAAGGTTTATGCCCAGATTA 429
Db 121 CTTTACACACCGCTTATGAATTTACTCCCGAGTCAGCAAGSITGTTGCTTGGATGTTG 180
QY 430 ATTTACCTGTGAAACATCATTCATATTTCTCCATGAACAGGAGTTCCTCTAGCTCTC 489
Db 181 ATCTACAGTGAACAGGCAATTTCTATATTTGCAAGAGCAGGAGTTTTCAGGCTCTC 240
QY 490 TCTGGGACTCAATTCAGAGTCAATTTGTTGCTCTCTTAGCCCATTTGCAATTTCTACTCA 549
Db 241 TTTGGGACTTCTGCAAGGCGCAATTTGTTGCTCTCTTAGGCTTGGATTTTATTTAA 300
QY 550 TATTGCGGAGCTAGAACTCATGGCTCGAATTTGACAGAGACAGCTTGAACACACA 609
Db 301 TTTTAAGAGAGCTGGGGAATCATGGATCCAATCTGACAGAGAGGAGCTTGAACACATA 360
QY 610 CTATATCTGCATGGAAGAGCTTAAAGCGCAAACTAATGGAAGAAATGATAGTCAAGT-- 667
Db 361 CCAATCAGCTTGGAAAGAGGAAATCGTATCTAAATAGACAGAACAAATGGCAATGGAA 420
QY 668 -GGAGCGCAACAGACTTAGTGTGATGCTGCAACCCCTTATGATCTCTGAGGAGCATTCAG 726
Db 421 CTGCAATTTTCAAGATGTTTATCCATGAGGSCCATATGATATCTGAAGATATTGCCA 480
QY 727 TAAAGCTTTTGAAGATGCGATTTTCTACAGTGCAGTTATTTATTCATCATCATCATG 786
Db 481 TGAAGATCTTGAAGAGGAGGTTTCAACTGTTCTATTTATCCATTTCTCTTGAAGATG 540
QY 787 GATCATTTCCGAGTTACTACATCTTGTCTTCACTTTTCAGGAATCTTAAATGCAATTTGTA 846
Db 541 CTTCAATTTCCAGTTACTACATCTTGTCTTCACTTTTCAGGAATCTTAAATGCAATTTGTA 600
QY 847 GATCATTTCAAAACTCAACTGGTAATTTGCTTCTTGAACCAACCAAGTGTCTTCAATTC 906
Db 601 GSTATTTTAGGCACCTGCTAGTGTCTTGTCTTGTCTTCACTTCACTTCAATCTGTGCAATAC 660
QY 907 GCTGCGGTTCTGGGTTCCGAAATTTGTTGATCTGGAACAGTGTCTTCAATTTGGCTATGTC 966
Db 661 CTGTGGGCACTGGGTTGCCAAATTTGGGGAATCAAAATCGCGCGCTCTTAGCAATGTTGA 720
QY 967 GACCTAATGCTCCTCCTAGTCTTGGCTTAAATGTTGGTTCAAGCTGGAGTAAAGTCAA 1026
Db 721 GACCAACCGCTTCTTGTGCTGAGCTTAAATTTATTTAGTTTCAAGCCCAAGTAAAGTCAA 780
QY 1027 TACCAATTTGGATGACAAAGCTCCCTGCTGACACTTCTCTAGAGTGCATCATCAG 1086
Db 781 TACCAATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1087 CCTAGCAAAAGACAGGCTTACACATGTTTGGCTGATGATGATGATGATGATGATGATGATGAT 1146
Db 841 CTTTGGCAAGAACAGAGCATATACATATTAATCTTGAAGAAATGATCTTGTTCATCAGG 900
QY 1147 CTTTGCAGCTTGGCAAGATGCAATATACACCTTTTGGATTTCTTAAAGCCAGAGATGCC 1206
Db 901 CTTTGCAGTTGGCCAGGAGCATATAGTCCCTATG---AGCTTAGAGTCAAGATGTC 957
QY 1207 AGATGCTCTCGGTTCTGATCTCTTGTCTGAAGTGTGAGGAGCTTGGCTTAATCTCTGGG 1266
Db 958 AGATGTTTGGGTTCTGATCTCTCTGATATAAGTGTGAGGAGCTTGGCAATCCAGGTG 1017
QY 1267 TCGGCGGCTGCTCTTGTGGAAGCTGGGAGCAAGAGTGTGGAGGCAATCATATCACTAA 1326

Db 1018 TCAGGGGGCTTGTATCTGCGAAGCTGGCGACGAGCGGTGAGAGCGCATGTTTCATTGA 1077

Qy 1327 GTGACATTTTCAAGTTCTT 1345

Db 1078 GTGACATATTCAGATTTCTT 1096

RESULT 9

US-10-424-599-79689
 ; Sequence 79689, Application US/10424599
 ; Publication No. US2004031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 79689
 ; LENGTH: 2315
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_42975C.1
 US-10-424-599-79689

Query Match 32.0%; Score 504.4; DB 13; Length 2315;
 Best Local Similarity 62.4%; Pred. No. 5.6e-142; Indels 48; Gaps 2;
 Matches 852; Conservative 0; Mismatches 466;

Qy 8 GGAACATTTCCGATGCTCTATCGAAGGCTGCCACCTGATTTTCAGGCTATTTGCAG 67

Db 581 GAACCATATGCTATGCTCCGATGGAGGATGCCCTGCTGTTTCAAGTTGTTGCAA 640

Qy 68 TCTGCTCTCGGATTCACGATACAAATTTTGTGGACGGGAATGGGGCATGATGA 127

Db 641 CTTAATGCTCGGGTTTCATCAGTACAAATTTAATGTAGATGGTGGTGGGCGACGATGA 700

Qy 128 GCGTCAACCTACCATATCTCGGGAGTTTGGCATAGTTTAAACACATTTTACTTGCACAAAGGA 187

Db 701 SCAGACCCATTTGTAATGGAGCTGTGGAGTAGTGAACACTATTTATATAGTGAGAGA 760

Qy 188 ATATAACAAATAAACACCTTATCAAGTCCAGACACCTCGGAAGCAGGATGAACATGGA 247

Db 761 ACCAGATATCTTACCTTCCATCTTAATATCTGAAACACCTGGTGCATCACACATGGAGGT 820

Qy 248 TGTGATAATGMAAATTTTCAACGTACGGTTACGTTGTGATGCGACCGCTTTCAGGAGG 307

Db 821 TGACACATG-----GAGC 835

Qy 308 TACTCTCAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGCTGTTTCTGAATATCT 367

Db 836 TAATCCAGGATGACTGTGCTGATCTGGAGTTTCTGCCACCGTATATCGTATTTCT 895

Qy 368 GAATTTGCATACATGCTATGATTTACTCCAGATTTCTGGCAAGTTATGCGCTAGACAT 427

Db 896 GTCCACGCACTTTCATATGATTTGCTTCCGAGTCAGGAAAGGTCAITGCTGCTGGATAT 955

Qy 428 TAATTTACCTGTGAGCAATCATTCATATCTCCATCAACAGAGGGATTCCTGTAGCTCC 487

Db 956 AAATTTACCATTTAAGCAAGCATTCATGTTCTTTTACCAAGGATATATCTATGCGCTCC 1015

Qy 488 TCTCTGAGCTCATTCAGAGGTCATTTTGTGCTCTCTAGCCCATTTGATTTTCATCT 547

Db 1016 TCTATGGATTTTTCAGAGCCAGCTGTTGGAGTTCTTAGTGAATGACTTCAITCT 1075

Qy 548 CATATTCGGAGCTAGAAATCTATGCTCGAACTTGTGACAGAGAGCGCTTGAACACA 607

Db 1076 AATATTGAAGAGCTGGGGAACCATAGTTTCAAAATTTGACTCAAGAAACAACTCGAGACTCA 1135

Qy 608 CACTATATCTGCATGGAAGAGCTGAGCGCAAACTAATGGAAGAAATCATAGTCACTG 667

Db 1136 TACTATAGCTGCTCGAAGAGAGAAATTTCAACATTCAGAACACTGTATAGTATGG 1195

Qy 668 GCGACCGCAAC---AGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGGGAGCATATGC 724

Db 1196 GGGATCATATCTCTGGCGATTTGTTTCATGCGCGACCCCATGAATGCTAAAGATGGC 1255

Qy 725 AGTAAAGCTTTTGCAGAAATGGCATTTCTACAGTGGCCAGTTATTTATTCATCATCAGAGA 784

Db 1256 TTTGAAGGTTTTCAGAAACAAAGGTGTCAACTGTCCCTATCATACATTTCTTCTCAGAGGA 1315

Qy 785 TGGATCATTTCCCGCAGTTATTTGCATCTTGCATCACTTTCTGGAATTTTGAATGTATTTG 844

Db 1316 TGGTTTCAATATCTCAACTGCTCATCTTGTCTCCCTATCCGGAATACTAAAGGATATG 1375

Qy 845 TAGATACTTCAAACAACTCAACTGGTAAATTTGCTTATTTCTGAAACCAACAGTGTCTCAT 904

Db 1376 CAGGCATTTTAAAGCACTCTCTGTGAGTTCTTTGCCCATTTCTTCAACTCCAGTGGCTTCAAT 1435

Qy 905 TCCGCTGGGTTCTCGGTTCCGAAATTTGGTATCTGAAACAGTCTCCATTTGGCTATGTT 964

Db 1436 ACCTTTGGGTACATGGATGCTTAGAGTTGGGGAACCAAAATGGTGGGCCACTAGCAATGTT 1495

Qy 965 GCGACCTAATGCTCACTTAGCTCTGCCCTTAAACATGTTGGTTCAAGCTGGAGTAAGCTC 1024

Db 1496 GATGCCAAGTGTCTCTCGTGTCTTGTCTATGTTGTTCAAGCCGAGTTAGCTC 1555

Qy 1025 AATACCAATTTGGATGACAAACGACTCCCTGCTGTGACACTTACTCTAGAGTGACATCAC 1084

Db 1556 AATACCGATTTGGATGCTAATGATTCGTTGCTTGACATTTATTCAGAGAGCGATATTAC 1615

Qy 1085 AGCCTAGCAAGAGCAAGGCTTACACACATGTTGGCTGGATGAGATGACCATTCACCA 1144

Db 1616 TGCATTTGGTAAAGATAAAGCTTATGCTAGATATCTCGATGAAATTTAGTATTCACCA 1675

Qy 1145 GGCTTTGCAGCTTGGCAAGATGCCAATAACACCTTTTGGATTTCTTAAACGCCAGAGATG 1204

Db 1676 GGCATTTGCTTTGGCAAGATGCACTTCTCTTATTTGGCTTTTACAATGCTCACAGATG 1735

Qy 1205 CCAGATGTGCTCGGCTGATCTCTGCTGAAAGGATGAGGAGCATGCGCTATCTCTGG 1264

Db 1736 TCACATGTTTGGATCTGATTCGCTTTTCAAAAGTGAATGAGCGGTTGGCTAATCTCTGG 1795

Qy 1265 GGTCCGCGGGTGTTCATTTGTTGAGCTGGAGCAAAAGTGTGAGGCGCATCATATCACT 1324

Db 1796 GGTAGGAGACTTGTGGCTGTGGAGGCTGGCAGCAAGCGTGTGAAGGGATTTTCTTT 1855

Qy 1325 AAGTGACATTTTCAAGTTCTTCTGCTGAGCTTGTGAGGAGAGAGGAGC 1370

Db 1856 AAGCGATGCTTTCAGATTTCTTTAGGCTAGTTGGGATATATTGGC 1901

RESULT 10

US-10-437-963-90730
 ; Sequence 90730, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966

US-10-424-599-101

```

Query Match      20.4%; Score 321.8; DB 13; Length 1225;
Best Local Similarity 66.4%; Pred. No. 1.5e-86;
Matches 541; Conservative 0; Mismatches 237; Indels 37; Gaps 4;

QY 540 TTCAATCTATATTTGGGGAGCTAGAACTCATGCTCGAATCTTACAGAGAGAGAGCTT 599
DB 3 TTGAACACATACCATCATCATCTGCTTGGAAATCATGGTCCAAATCTAACAGAGAGAGCTT 62
QY 600 GAAACACACATATCTGATGTAAGAGAGAGCTAAGCGGCAAACTTAATGGAAGAAATGAT 659
DB 63 GAAACACATACATATCATGCTTGGAAAGAGAGAAATGGAAGAGAGAGAGAGAGAGAG 105
QY 660 AGTCAGTGGGACCGCAACAGCATCTAGTCATGCCACCCCTTATGAGTCCCTTGAGGAG 719
DB 106 -----GATTTACACATATGTTTATCCGTGAGGAGGACATAGATAATTTGAAGAG 155
QY 720 ATTCAGTAAAGCTTTTGCAGAAATGGCAATTTCTACAGTGCAGTATTTATTTATCATCA 779
DB 156 ATTCAGTAAAGCTTTTGCAGAAATGGCAATTTCTACAGTGCAGTATTTATTTATCATCA 209
QY 780 TCAGATGATCATTTCCGAGTATTTGATCTTGTGATCACTTTCTGGAATTTGAAATGT 839
DB 210 GAAGATGCTCATTTCCAGCTACTACATCTTGCTTCACTTTCAGGAAATCTTAAATGC 269
QY 840 ATTTGTAGATCTTCAAAACTCAACTGGTAAATTTGCTTATTTCTGAACTTGAAGTGC 899
DB 270 ATTTGTAGATCTTCAAAACTCAACTGGTAAATTTGCTTATTTCTGAACTTGAAGTGC 329
QY 900 TCCATTCGCTGGTTCCTGGTTCGAAATTTGATGATCTGAACTGCTGCTTCAATGGCT 959
DB 330 GCAATTCCTGTGGGACGCTGGTGGTCCCAAAATTTGGGAGTCAAAATTCGCGGCTCTAG 389
QY 960 ATGTTGGACCTAATGCTCTAGCTGCTGCTTCAATGCTTGGTCAAGCTGAGTA 1019
DB 390 ATGTTGGACCTAATGCTCTAGCTGCTGCTTCAATGCTTGGTCAAGCTGAGTA 449
QY 1020 AGCTCAATACCAATTTGGATGACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
DB 450 AGTTCAATACCAATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 509
QY 1079 CATCACGCTTAGCAAGAGCAAGTCTACACATGCTTGGCTGCTGCTGCTGCTGCTGCT 1138
DB 510 TATAACAGCTTTGGCAAGAGCAGAACTTATACCCATATTAATCTTGATGAAATGACT 569
QY 1139 TCACAGGCTTTGACAGCTTGGCAAGAGTCCAAATACACCTTTTGGATTTCTTAAAGGCA 1198
DB 570 TCATCAGGCTTTGCAATTTGGGCGAGATTTCTTATAATCTTATG---AGCTGAGTTGCA 626
QY 1199 GAGATGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1258
DB 627 AGATGTCAGATGCTTTTGGCAAGTGAATCTCTGCAATAAGTGTGAGAACTTTGGCAA 686
QY 1259 TCTCGGGTGGCGGGTCTTCAATTTGGAAGCTGGGCAAAAGTGTGAGGAGGATCAT 1318
DB 687 TCCAGTGTGAGCGGCTTGTAAATTTGGAAGCTGGCAGCAAGCGGTAGAGAGGATCAT 746
QY 1319 ATCACTAAGTGAATTTTCACTTCTTCTGCTGAGCT 1353
DB 747 AGCATTTGATGACATATTTCACTTCTTCTGCTTGGT 781

```

RESULT 13

```

US-10-425-114-30669
; Sequence 30669, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

```

```

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30669
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMELB73061F11_F11
US-10-425-114-30669

```

```

Query Match      19.7%; Score 310; DB 13; Length 721;
Best Local Similarity 95.9%; Pred. No. 4.1e-83;
Matches 353; Conservative 0; Mismatches 5; Indels 10; Gaps 3;

QY 1143 CAGGCTTTGAGCTTGGACAGATGCCAATACACCTTTTGGATTTCTTAAACGCCAGACA 1202
DB 288 CAGGCTTTGAGCTTGGACAGATGCCAATACACCTTTTGGATTTCTTAAACGCCAGACA 347
QY 1203 TGCCAGATGTGCTCCGCTCTCATCTTCTGCTGAAGTGAAGCTGAGCGACTGGCTAATCT 1262
DB 348 TGCCAGATGTGCTCCGCTCTCATCTTCTGCTGAAGTGAAGCTGAGCGACTGGCTAATCT 407
QY 1263 GGGGTGGCGGGGTGCTTCAATGTGAAGCTGGAGCAAAACGCTGGAGGCGATCATATCA 1322
DB 408 GGGGTGGCGGGGTGCTTCAATGTGAAGCTGGAGCAAAACGCTGGAGGCGATCATATCA 467
QY 1323 CTAAGTGACATTTTCAAGTTCCTGCTGAGCTGAGGAGAGAGGAGGAGGAGGAGGAGT 1382
DB 468 CTAAGTGACATTTTCAAGTTCCTGCTGAGCTGAGGAGAGAGGAGGAGGAGGAGGAGT 526
QY 1383 TTTTAAACAGTAGCAACCTCTGGGCTTTGGGCTCTTGGGCTCTTGGGCTATGCTCAG 1442
DB 527 TTTTAAACAGTAGCAACCTCTGGGCTTTGGGCTCTTGGGCTATGCTCAG---582
QY 1443 GAGCTCAGGCTTCAAGTCTGTAAGTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1501
DB 583 ----TCAGGCTTCAAGTCTGTAAGTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 638
QY 1502 GCCCTCCC 1509
DB 639 GCCCTCCC 646

```

RESULT 14

```

US-09-294-093B-135
; Sequence 135, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700342215H1
US-09-294-093B-135

```

Query Match 16.9%; Score 266.2; DB 9; Length 282;
Best Local Similarity 98.6%; Pred. No. 4.7e-70;
Matches 279; Conservative 0; Mismatches 3; Indels 1; Gaps 2;
QY 86 CGAGTACAAATCTTTGGACGGGGAATGGCGGATGATGAGCGTCAACCTTACCATATC 145
DB 1 CGAGTACAAATCTTTGGACGGGGAATGGCGGATGATGAGCGTCAACCTTACCATATC 60
QY 146 TGGGAGGTTTGGCAGATTTAAACACACTTTACTTGAACAGGGAATATAACCAATAAACAC 205
DB 61 TGGGAGGTTTGGCAGATTTAAACACACTTTACTTGAACAGGGAATATAACCAATAAACAC 120
QY 206 CTTATCAAGTCCCAAGCACACCTTGAAGAGGAGTGAACATGAGTGTGGTAATGAATAATTT 265
DB 121 CTTATCAAGTCCCAAGCACACCTTGAAGAGGAGTGAACATGAGTGTGGTAATGAATAATTT 180
QY 266 TCAAGCTAGCTTACGTTCTGATGCGACCGTTTCAGAGGTPACTCTGAGAGTTTCAGA 325
DB 181 TCAAGC-GGGGTTACGTTCTGATGCGTACCGTTTCAGAGGTPACTCTGAGAGTTTCAGA 239
QY 326 GGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTG 368
DB 240 GGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTG 282
RESULT 15
US-10-437-963-66100
; Sequence 66100, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbarov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 66100
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67084C.1
US-10-437-963-66100
Query Match 15.3%; Score 241.6; DB 17; Length 1402;
Best Local Similarity 54.0%; Pred. No. 4.3e-62;
Matches 552; Conservative 0; Mismatches 449; Indels 21; Gaps 2;
QY 321 TCAGAGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTGAATTCGATACA 380
DB 343 TCAGGCGTGCAATACAAATTTCCGGCATCGAGTCTCTGAGATATTTATACATAATACC 402
QY 381 TGCTATATTTACTCCAGATTTGGGAGGTTATTCCTGATAGATTAATTTACCTGTG 440
DB 403 ATATATGATGTGTTCTGTTTCTAGCAAGATAGCAGTTTGGATGCTGCTCTCTCTGTC 462
QY 441 AAGCAATCATTCATATTTCTCCATGCAACAGGGGATTCCTGTAGTCTCTCTCTGGACTCA 500
DB 463 AAGCAGCTTTAAATCATCATGATGAGGCTCTTTCTTGGTTCCTCTTTGGATGAT 522
QY 501 TTCAGAGCTCAATTTGTTGTTCTCTTGTAGCCCATTTGGAATTTCACTCATATTTGGGAG 560
DB 523 CAGCAGCAAAACCGTAAACAGGATGCTAAACCGCATCCGATTTGTTATTAATTTTGAGAAAG 582
QY 561 CTAGAACTCATGCTCGAATTTGACAGAGGAGCTTTGAACACACACACTATATCTGCA 620

DB 583 CTGACAGAGAAACATTCGAACCTCTTGCCATGAGGAACCTTGAATGATTCAGTGTCTGCT 642
QY 621 TGGAAAGAGGCTTAAGCGCAAACTAATGGAAGAAATGATAGTCACTGAGTGGCGACCGACAG 680
DB 643 TGGAAAGAGCAAAATTTGCAGTTTATGCGGGGACCTGATGTTGCTGCCATTCAGAGAGG 702
QY 681 CATCTAGTGCATGCCACCCCTTATGATGCTTGGAGGACATTCAGTAAGCTTTTGCAA 740
DB 703 CCATTAATTCATGTTAAGGATTCAGATAATTTAAGGATGTCGATGTTGCTATATATAGA 762
QY 741 AATGGCAATTTACAGTGCAGTATTTATTCATCATCATCATCATCATCATCATTCCTCCG 800
DB 763 AATGAATATCTTCAGTTCCTTAACTTTAAGCCCTCAACGGATTCATCAGGATGCCT 819
QY 801 TTATTCATCTTTCATCATCTTCTGGAATTTTGAATTTTGAATTTTGAATTTTGAATTTT 860
DB 820 TTGCTTGGTCTTTCGAACCCCTTCCAGGATTTGGAATTTTATTTGCTCAAGCTGCAAGAA 879
QY 861 TCAACTGTAATTTGCTTATCTGAACCAACAGTGTGCTCCATTCGGCTGGGTTCTCTGG 920
DB 880 CAGCCTGAAAGGGTACTCATTTCTGCAAAACAGATTCAGTATGCTATTTGCTATGATGG 939
QY 921 GTTCCGAAATTTGCTGATCTGAACAGTGTGCTCCATTCGGCTATGTTGGACCTTAATGCTCA 980
DB 940 TCACCATATCTGCGAAGGCAAGCAATAGACAGCTTAGAATTCGCGACCAAGCACTCCT 999
QY 981 CTTAGCTCTGCCCTTAACTGTTTCAAGCTGAGTGAAGTCAATACCAATTTGTTGAT 1040
DB 1000 CTAAATTCATGCTGGAATTTGCTGCTTGAAGATAGTAAGTCAATTTCTATATAGTTGAC 1059
QY 1041 GACACGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
DB 1060 GATAAATGGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119
QY 1101 AAGTCTTACACACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160
DB 1120 GATGCTTACACTCGTATTGAGCTTGAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1179
QY 1161 CAAGATCCCAATACACCTTTTGGATTTCTTTAAAGCCAGAGATGCGAGATGCTGCTCGG 1220
DB 1180 TACCAGGTGAA-----TGGCCGAAAGACACTGTCTATACCTGCTTGAGC 1221
QY 1221 TCTGATCCTTTGCTGAAGGTGATGAGCGACTGCTGCTTAATCCTGCTGCTGCTGCTGCTGCT 1280
DB 1222 ACTGATACCTTCTGAGGTTTGGAGCAATTTGTCAGTCCAGGGTGGCGGAGTCTGCTC 1281
QY 1281 ATTGTGAAGCTGGGAGCAACAGTGTGAGGGCATCATATCATTAAGTGAATTTTCAAG 1340
DB 1282 GTTATTGAACCAAGGAGCAGATTTGTGCAAGGAATTAATCTCATTTGAGGAGCGCATTTACA 1341
QY 1341 TT 1342
DB 1342 TT 1343

Search completed: July 9, 2004, 08:27:50
Job time : 783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:46:38 ; Search time 59 Seconds
(without alignments)
2159.812 Million cell updates/sec

Title: US-09-857-525C-2
Perfect score: 2342
Sequence: 1 TREHLPMSPIEGPTVFQA.....SKRVEGIISLSDIFKFLLSL 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	451	3	AAY96783 Partial Z
2	2085	89.0	493	3	AAY96784 Rice suc
3	2019	86.2	442	3	AAY96789 Wheat suc
4	1581	67.5	492	3	AAY96787 Soybean s
5	1525.5	65.1	482	3	AAY96786 Soybean s
6	1214	51.8	368	5	AB90919 Herbicida
7	392	16.7	189	3	AAY96785 Parital r
8	373	15.9	328	2	AAW88438 Disease a
9	373	15.9	328	4	AAW93432 Human ins
10	373	15.9	328	5	AAO18496 Human ins
11	373	15.9	352	5	AAO18495 Human ins
12	369	15.8	77	3	AAY96793 Wheat suc
13	367	15.7	331	2	AAW29817 Mammalian
14	367	15.7	331	5	AAO18497 Human ins
15	367	15.7	331	6	ABU04258 Human exp
16	367	15.7	331	6	ABU04261 Human exp
17	367	15.7	331	6	ABU04262 Human exp
18	367	15.7	331	6	ABU04257 Human exp
19	367	15.7	331	7	ABU04263 Human pro
20	367	15.7	344	3	ABU04259 Human pan
21	367	15.7	344	6	ABU04259 Human exp
22	361.5	15.4	634	4	ABW6245 Drosophil
23	361.5	15.4	634	5	ABG70033 Larval vi
24	355.5	15.2	330	7	ADE62621 Rat Prote
25	340.5	14.5	305	4	AAE00329 Human Prk

26	340.5	14.5	305	4	AAE00328 Human Prk
27	340	14.5	464	4	AAE00223 Human AMP
28	340	14.5	489	4	AAW47679 PRKAG3. 1
29	340	14.5	489	5	ABP43929 AMP activ
30	340	14.5	489	6	AAE32034 Human Kin
31	339.5	14.5	305	4	AAE00221 Human AMP
32	335.5	14.3	305	4	AAE00226 Sus scrof
33	335.5	14.3	325	4	AAE00225 Sus scrof
34	335.5	14.3	333	5	ABB06101 Human NS
35	335.5	14.3	464	5	AAE22987 Pig PRKAG
36	335.5	14.3	464	5	AAE22988 Pig PRKAG
37	334.5	14.3	305	4	AAE00220 Pig AMPK
38	334.5	14.3	464	4	AAE00222 Pig AMPK
39	334.5	14.3	464	5	AAE22984 Pig wild-
40	334.5	14.3	464	5	AAE22985 Pig PRKAG
41	334.5	14.3	464	5	AAE22986 Pig PRKAG
42	334.5	14.3	514	4	AAE00224 Sus scrof
43	329	14.0	489	7	AAE38410 Human Prk
44	303.5	13.0	322	6	ABR52942 Protein s
45	302	12.9	488	4	ABG20078 Novel hum

ALIGNMENTS

RESULT 1

AAAY96783
ID AAY96783 standard; protein; 451 AA.

AC AAY96783;

DT 26-SEP-2000 (first entry)

DE Partial Z. may's sucrose non-fermenting 4 protein.

KW Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
KW carbon catabolite repression; development; nitrogen partitioning.

XX Zea mays.

FN WO200036116-A2.

PD 22-JUN-2000.

PF 15-DEC-1999; 99WO-US029825.

PR 16-DEC-1998; 98US-0112564P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Allen SM, Heppard EP, Miao G, Weng Z;

XX WPI; 2000-431594/37.

XX N-PSDB; AAA51347.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
involved in carbon catabolite repression in plants and seeds, useful for
controlling carbon and nitrogen partitioning pathways during plant growth
and development.

XX Claim 10; Page 31-33; 48pp; English.

XX This is a partial maize (Zea mays) sucrose non-fermenting 4 (SNF4)
protein which is involved in carbon catabolite repression in plants and
seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
csk1c.pk001.c15 prepared from corn unpollinated developing silk 24 hours
after emergence. The polynucleotides are used in plants to control carbon
and nitrogen partitioning pathways during plant growth and development.
XX The catabolite repression proteins would facilitate studies for better
understanding the mechanism of catabolite repression in plants and could
provide genetic tools to enhance or otherwise alter the accumulation of
carbohydrates, lipids and proteins during plant growth and development

XX PF 15-DEC-1999; 99WO-US029825.
 XX PR 16-DEC-1998; 98US-0112564P.
 XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX DR MPI; 2000-431594/37.
 XX DR N-PSDB; AAAS1353.
 XX XX
 PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX XX
 ES Claim 24; Page 44-45; 48pp; English.
 XX XX
 CC This protein is a wheat sucrose non-fermenting 4 (SNF4) protein which is
 CC involved in carbon catabolite repression in plants and seeds. The cDNA
 CC was isolated, based on similarity to SNF4 proteins from Arabidopsis
 CC thaliana and Saccharomyces cerevisiae, from library wll prepared from
 CC wheat leaf from 7 day old seedling. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX XX
 SQ Sequence 442 AA;

Query Match 86.2%; Score 2019; DB 3; Length 442;
 Best Local Similarity 86.5%; Pred. No. 3.3e-196;
 Matches 384; Conservative 33; Mismatches 25; Indels 2; Gaps 1;
 QY 7 MSPIEGCTVFOAICSLSPGHIHYKFFVGDGWRHDERQPTISGEFGVNTLYLTREYNQI 66
 DB 1 MSPVEGCTVFOAICNLPPIYQYKENVGQWRHDEGQPTTIGEYGVNTLYLTREFDHI 60
 QY 67 NTLSPSTFGSRMMDVNEFNQRTVTLSDGTVSEGLRVSEAAIQISCRVSEYL 126
 DB 61 NTVLSPSTFESR--MDVDSFSFORMGSLDQALQEGSPRISEAAIQISCRVAEYLNHT 118
 QY 127 CYDLLPDSGKVIALDINLVPKQSFHILHEQGIQVAPLWDSFRGQVGLLSPDLFILRE 186
 DB 119 GYDLLPDSGKVIALDINLVPKQSFHILHEQGIQVAPLWDSFRGQVGLLSPDLFILRE 178
 QY 187 LETHGSLNTEEOLETHITISAWKEAKQTNQGRNDSQWRPOOHLVHATPYESLRDIAVKLLQ 246
 DB 179 LETHGSLNTEEOLETHITISAWKEAKQTYGRNDQGRNDSQWRPOOHLVHATPYESLRGIAKMLE 238
 QY 247 NGISTVPIYSSSSDGSFPOLLHLASLSGILKICRYFKNSTGNLPILNQVCSIPLGSW 306
 DB 239 TGISTVPIYSSSSDGSFPOLLHLASLSGILKICRYFKNSTGNLPILNQVCSIPLGFW 298
 QY 307 VPKIGDLNSRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLDITYSRSDITALAKD 366
 DB 299 VPKIGEPNGHPLAMLRPNSTLSSALNMLVQAGVSSIPVDDNDSLDITYSRSDITALAKD 358
 QY 367 KYTVHRLDEMTIHOALQIQDANTPFQFNQRCOMCLRSDFLLKXWERLANPGVRRVF 426
 DB 359 KYTVHRLDEMTIHOALQIQDANSFGLFNQRCOMCLQSDFLKXWERLANPGVRRVF 418
 QY 427 IVEAGSKRVEGIISLIDFKLLS 450
 DB 419 IVEAGSKRVEGISLIDFKLLS 442
 RESULT 4
 AAY96787
 ID AAY96787 standard; protein; 492 AA.

XX AC AAY96787;
 XX DT 26-SEP-2000 (first entry)
 XX DE Soybean sucrose non-fermenting 4 protein from clone sfl1.pk0004.b.
 XX KW Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX KW carbon catabolite repression; development; nitrogen partitioning.
 XX OS Glycine max.
 XX XX
 PN WO2000036116-A2.
 PD 22-JUN-2000.
 PF 15-DEC-1999; 99WO-US029825.
 PR 16-DEC-1998; 98US-0112564P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PA Allen SM, Heppard EP, Miao G, Weng Z;
 XX DR MPI; 2000-431594/37.
 XX DR N-PSDB; AAAS1351.
 XX XX
 PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX XX
 PS Claim 10; Page 40-41; 48pp; English.
 XX XX
 CC This is a soybean (Glycine max) sucrose non-fermenting 4 (SNF4) protein
 CC which is involved in carbon catabolite repression in plants and seeds.
 CC The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from library sfl1
 CC prepared from soybean immature flower. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX XX
 SQ Sequence 492 AA;

Query Match 67.5%; Score 1581; DB 3; Length 492;
 Best Local Similarity 67.6%; Pred. No. 1.6e-151;
 Matches 305; Conservative 61; Mismatches 75; Indels 10; Gaps 4;
 QY 3 EHLMPSPGCTVFOAICSLSPGHIHYKFFVGDGWRHDERQPTISGEFGVNTLYLTRE 62
 DB 46 ELLMPSPGCTVFOAICNLPPIYQYKENVGQWRHDEGQPTTIGEYGVNTLYLTRE 105
 QY 63 YNOITLSPSTFGSRMMDVNEFNQRTVTLSDGTVSEGLRVSEAAIQISCRVSEYL 122
 DB 106 PNYMPVL--PPDVASGNSMDVDNDAPFRMARLTDGTLSEVLPRISDTQVQISQRISAF 163
 QY 123 NLHTCYDLLPDSGKVIALDINLVPKQSFHILHEQGIQVAPLWDSFRGQVGLLSPDLFIL 182
 DB 164 SSHTAYELLPEGSKVVALDVLVQAQPHILHEQGVFMALDFCKGQGVGLSASEFIL 223
 QY 183 ILRELATHGSLNTEEOLETHITISAWKEAK----RQTNQGRNDSQWRPOOHLVHATPYESLR 238
 DB 224 ILRELATHGSLNTEEOLETHITISAWKEAKSYLNQNGHGTAFSR---CFIHAGEYDNLK 280
 QY 239 DIAVKLLQNGISTVPIYSSSSDGSFPOLLHLASLSGILKICRYFKNSTGNLPILNQV 298
 DB 281 DIAVKLLQNGISTVPIYSSSSDGSFPOLLHLASLSGILKICRYFKNSTGNLPILNQV 340
 QY 299 CSIPIGSWVPKIGDLNSRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLDITYSR 358

```

Db 341 CAIPVGVWPKIGSNRRPLAMLRRTASLASAKLLVQAQVSSIPVDDNDSDLDIYCRS 400
Qy 359 DITALAKDKVYTHVRLDEMTIHQALQLGODANTPFQFFNGORCOMCLRSDDPLKQWERLA 418
Db 401 DITALAKNRAYTHINLDEMTVHOALQLGODAVSPYE-LRSQCQCXCLRSDDPLKQWERLA 459
Qy 419 NPGVRRVIVEAGSKRVEGIISLSDIFKELL 449
Db 460 NPGVRRVIVEAGSKRVEGIISLSDIFKELL 490

RESULT 5
ID AAY96786 standard; protein; 482 AA.
XX AAY96786;
XX AAY96786;
DT 26-SEP-2000 (first entry)
DE Soybean sucrose non-fermenting 4 protein from clone ses4d.pk0040.b1.
XX
KW Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
KW carbon catabolite repression; development; nitrogen partitioning.
XX
OS Glycine max.
XX
XX WO200036116-A2.
XX
XX 22-JUN-2000.
XX
XX 15-DEC-1999; 99WO-US029825.
XX
XX 16-DEC-1998; 98US-0112564P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Heppard EP, Miao G, Weng Z;
XX
XX WPI; 2000-431594/37.
XX
XX N-PSDB; AAA51350.
XX
XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
XX involved in carbon catabolite repression in plants and seeds, useful for
XX controlling carbon and nitrogen partitioning pathways during plant growth
XX and development.
XX
XX Claim 10; Page 37-39; 48pp; English.
XX
XX This is a soybean (Glycine max) sucrose non-fermenting 4 (SNF4) protein
XX which is involved in carbon catabolite repression in plants and seeds.
XX The cDNA was isolated, based on similarity to SNF4 proteins from
XX Arabidopsis thaliana and Saccharomyces cerevisiae, from library ses4d
XX prepared from soybean embryogenic suspension 4 days after subculture. The
XX polynucleotides are used in plants to control carbon and nitrogen
XX partitioning pathways during plant growth and development. The catabolite
XX repression proteins would facilitate studies for better understanding the
XX mechanism of catabolite repression in plants and could provide genetic
XX tools to enhance or otherwise alter the accumulation of carbohydrates,
XX lipids and proteins during plant growth and development
XX
XX Sequence 482 AA;

Query Match 65.1%; Score 1525.5; DB 3; Length 482;
Best Local Similarity 67.5%; Pred. No. 7.1e-146;
Matches 303; Conservative 57; Mismatches 72; Indels 17; Gaps 6;

Qy 3 EHLPMSPFGCPTVFOAICSLSPGIEHYKFFVGVGWRHDERQPTISGFGIVNTLYLTRE 62
Db 45 ELLQMSFVGVGCTPVFQVHSLVFGHHQYKFFVGVGWRHDDLQPCSEGEYGVNTVSLATD 104
Qy 63 YNQINTLSPFPGSNMMDVNMENFQRTVWTLSDGTVSEGL-RVSEMAIQISRCRVSEY 121

```

```

Db 105 PNILPVLTPDIVSGS--NMDVDNEAFRRMVRULTDGTLSNVLLPRISDVDTQTSQRISAF 162
Qy 122 LNLHTCYDLLPDSGVKIALDINLPVKQSPHILHEQGIPIVAPLWDSFRGQFVGLSLPDLFI 181
Db 163 LMSSTAYELLPSGKVTLDVLPVKQAFPHILHEQGIPIAPLWDICKGQFVGVLSDLFI 222
Qy 182 LILRELETHGSLNTEQLETHYTSIAWKEAKROTNGRNDSSQWRP-QOHLVHATPYESLRDI 240
Db 223 LILRELGNHGSNLTEELETHYTSIAWKGK-----WGFQTCF-RAGPYDNLKEI 272
Qy 241 AVKLLONGISLTVPIYSSSDGSPQLLHLASLSGLKICICRYFKNSTGNLPIKQPVCS 300
Db 273 AVKILQHGISTVPIIH--SEDSFPQLLHLASLSGLKICICRYFRNCSSSLPIQLPICA 330
Qy 301 IPLGSWVPKIGDLNSRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSDLDIYCRSDI 360
Db 331 IPVGTWVPKIGESNRPPLAMLRPNASLTSALNMLVQAGVSSIPVDDSDSDLDIYCRSDI 390
Qy 361 TALAKDKVYTHVRLDEMTIHQALQLGODANTPFQFFNGORCOMCLRSDDPLKQWERLAMP 420
Db 391 TALAKDRYTHINLDEMTVHOALQLGQDSYNTYB-LSCQRCQCLRTDLSLHKVMERLASE 449
Qy 421 GVRVRIIVEAGSKRVEGIISLSDIFKELL 449
Db 450 GVRVRIIVEAGSKRVEGIISLSDIFNFFL 478

RESULT 6
ABB90919
ID ABB90919 standard; protein; 368 AA.
XX ABB90919;
XX ABB90919;
DT 31-MAY-2002 (first entry)
DE Herbicidally active polypeptide SEQ ID NO 130.
XX
XX Herbicidally active polypeptide SEQ ID NO 130.
XX
XX Herbicidally active polypeptide SEQ ID NO 130.
XX
XX Arabidopsis thaliana.
XX
XX WO200210210-A2.
XX
XX 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP009892.
XX
XX 28-AUG-2001; 2001WO-EP009892.
XX
XX (FARB) BAYER AG.
XX
XX Tietjen K, Weidler M;
XX
XX WPI; 2002-269010/31.
XX
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX
XX Claim 5; SEQ ID NO 130; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides
XX
XX Sequence 368 AA;

```



```
PR 12-JAN-2001; 2001US-0261457P.
PR 12-JAN-2001; 2001US-0261458P.
PR 12-JAN-2001; 2001US-0261459P.
PR 12-JAN-2001; 2001US-0261461P.
PR 12-JAN-2001; 2001US-0261518P.
PR 12-JAN-2001; 2001US-0261531P.
PR 12-JAN-2001; 2001US-0261532P.
PR 12-JAN-2001; 2001US-0261589P.
PR 12-JAN-2001; 2001US-0261590P.
PR 12-JAN-2001; 2001US-0261634P.
PR 12-JAN-2001; 2001US-0261695P.
PR 12-JAN-2001; 2001US-0261697P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Seidel-Dugan C, Ferguson KC, Kidd T;
XX WPI; 2002-599664/64.
XX N-PSDB; AAL48614.
XX
XX Identifying an insulin receptor signaling modulator, useful as drug
XX targets for treating diabetes or metabolic disorders, comprises
XX contacting an assay system comprising insulin receptor signaling
XX modifiers with a test agent.
XX
XX Disclosure; Page 49-50; 232pp; English.
XX
XX The present invention relates to a method of identifying a candidate
XX insulin receptor (INR) signaling modulating agent, involving contacting
XX an assay system comprising an insulin receptor signaling modifier (ISM)
XX polypeptide or nucleic acid with a test agent, and detecting a test agent
XX biased activity of the assay system. The method is useful for
XX identifying candidate INR signaling modulating agents. ISM genes may be
XX used as drug targets for treatment of disorders related to INR signaling
XX such as diabetes or metabolic syndrome. ISM nucleic acids and
XX polypeptides are useful for identifying and testing agents that modulate
XX ISM function and for other applications related to the involvement of ISM
XX in INR signaling, and for identifying subjects having a predisposition to
XX such diseases associated with INR signaling. The present sequence is an
XX ISM protein described in the exemplification of the invention
XX
XX Sequence 352 AA;
XX
XX Query Match 15.8%; Score 373; DB 5; Length 352;
XX Best Local Similarity 29.8%; Pred. No. 1.1e-28;
XX Matches 104; Conservative 75; Mismatches 128; Indels 42; Gaps 13;
XX
XX 104 LRVSEAAIQISRCRV-SEYLNHRCYDLPDPSGKVIADINLPVKQSFHILHEQGIFVAP 162
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 29 LEFEDEAVESGCVYMRFSHKCYDIVPTSSKLVVFDTLQVKKAFFALVANGVRAAP 88
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 163 LWSDFRCQFVGLISPLDPIILIRELTHGNSLNE-EQLETHITISAWKEAKROTGRNDSQ 221
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 89 LWESKKQSPFGMLTITOFINILHRY--YKSPMVQIYELEEKHETWRELXYLQ-----ET 140
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 222 WRPQOHLVHATPYESLRDIAVKLLQNGISTVPVIYSSSDGSPFQLHLASLSGLLKIC 281
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 141 FKP---LVNISPDSASLDVAVYSLINKKHRLPVIDPISGNA-----LYLTHKILKFL- 191
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 282 RYFKNSTGNILNPQVCSIPFGSVKPKIGDLSRPLAMLPNASSLSALMMLVOAGVSS 341
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 192 QLFMSDMPKPAFMKQNLDELIGTV-----HNIAFIHPTPIIKALNIFVERRISA 242
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 342 IPIVDDNSLDDTSRSDITALAKQVYTHVRLDEMTHTQALQGDANTFFGFNGQRC 401
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 243 LPVWDESGKVDIYKEDVINAAEKTYNNL---DITVTQALQHSQ-----YFEG--V 291
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 402 QMCLRSDPLKLMERLANPGVRRYFIV-EAGSKRVEGIIISLSDIFKFL 449
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 292 VKCNKLEILTIVDRIVRAEVRHVLVVNEDS--IVGLISLSDILQAL 338
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX
XX RESULT 12
```

```
AA96793
ID AAY96793 standard; protein; 77 AA.
XX
XX AAY96793;
XX
XX 26-SEP-2000 (first entry)
XX
XX Wheat sucrose non-fermenting 4 protein from clone wreln.pk0143.e2.
XX
XX Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
XX carbon catabolite repression; development; nitrogen partitioning.
XX
XX Triticum aestivum.
XX
XX WO200036116-A2.
XX
XX 22-JUN-2000.
XX
XX 15-DEC-1999; 99WO-US029825.
XX
XX 16-DEC-1998; 98US-0112564P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Heppard EP, Miao G, Weng Z;
XX
XX WPI; 2000-431594/37.
XX N-PSDB; AAA51357.
XX
XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
XX involved in carbon catabolite repression in plants and seeds, useful for
XX controlling carbon and nitrogen partitioning pathways during plant growth
XX and development.
XX
XX Claim 23; Page 48; 48pp; English.
XX
XX This is a partial wheat sucrose non-fermenting 4 (SNF4) protein which is
XX involved in carbon catabolite repression in plants and seeds. The cDNA
XX was isolated, based on similarity to SNF4 proteins from Arabidopsis
XX thaliana and Saccharomyces cerevisiae. The polynucleotides are used in
XX plants to control carbon and nitrogen partitioning pathways during plant
XX growth and development. The catabolite repression proteins would
XX facilitate studies for better understanding the mechanism of catabolite
XX repression in plants and could provide genetic tools to enhance or
XX otherwise alter the accumulation of carbohydrates, lipids and proteins
XX during plant growth and development
XX
XX Sequence 77 AA;
XX
XX Query Match 15.8%; Score 369; DB 3; Length 77;
XX Best Local Similarity 90.9%; Pred. No. 2.3e-29;
XX Matches 70; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
XX
XX 228 LVHATPYESLRDIAVKLLQNGISTVPVIYSSSDGSPFQLHLASLSGLKICRYFKNS 287
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 1 LVHATPYESLRGIAMKILETGISTVPVIYSSSDGSPFQLHLASLSGLKICRYFKNS 60
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 288 TGNLPILNPQVCSIPLG 304
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 61 TGSILNPQVCSIPLG 77
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX
XX RESULT 13
XX AAW29817
XX ID AAW29817 standard; protein; 331 AA.
XX
XX AAW29817;
XX
XX 16-MAR-1998 (first entry)
XX
XX Mammalian AMP-gamma subunit protein.
XX
XX 5'-AMP activated protein kinase; AMPK; non-catalytic subunit;
```

KW protein phosphorylation; cholesterol; fatty acid; gamma subunit;
 XX hormone sensitive lipase; HSL; ss.
 OS Mammalia.

XX WO9725341-A1.
 XX 17-JUL-1997.

XX 07-JAN-1997; 97WO-US000270.
 XX 08-JAN-1996; 96AU-00007450.

XX (SVIN-) ST VINCENTS INST MEDICAL RES.
 XX (DART-) DARTMOUTH COLLEGE.

XX Kemp BE, Stapleton DL, Mitchelhill KI, Witters LA;
 XX WPI; 1997-372811/34.

XX New isolated 5'-AMP-activated protein kinase subunit(s) - used to develop
 XX products for treating e.g. hyper-cholesterolaemia, obesity, hypoxia,
 XX ischaemia, nutrition disorders or diabetes mellitus.

XX Claim 21; Page 54; 63pp; English.

XX This polynucleotide sequence encodes a mammalian 5'-AMP-activated protein
 XX kinase (AMPK) non-catalytic gamma subunit. AMPK polypeptides can be used
 XX to identify compounds which regulate the action of kinases. Such
 XX fragments can be used to reduce biosynthesis of cholesterol and fatty
 XX acids. They may also be used to inhibit the release of these molecules
 XX from intracellular stores by hormone sensitive lipase (HSL). They may
 XX also be used to reduce cellular malonyl CoA levels and promote the beta-
 XX oxidation of fatty acids by mitochondria. AMPK-alpha fragments could be
 XX used in the treatment of e.g. hypercholesterolaemia, hyperlipidaemia,
 XX obesity, clinical syndromes associated with hypoxia or ischaemia (e.g.
 XX myocardial infarction) disorders of nutrition and diabetes mellitus

XX Sequence 331 AA;

Query Match 15.7%; Score 367; DB 2; Length 331;

Best Local Similarity 29.1%; Pred. No. 4.1e-28;

Matches 111; Conservative 74; Mismatches 133; Indels 64; Gaps 14;

QY 69 LGSPTPGSRMNDVNFQRTVTLSLSDGTGTVSGTLRVSEAAIQISRCRVSEVNLHTCY 128
 DB 5 ISSDSPA-----VENEHPQETPE-SNNSV-----YTSPKSHRCY 39

QY 129 DLLPSGKVALDINLPVQSFHILHEQGIPIVAPLWDSFGQVGLLSPDLFILILRELE 188
 DB 40 DLIPYSSKLWFDTSQVKKAFALVTNGVRAAPLWDSKQSFVGMLTITDFINILHRY - 98

QY 189 THGSNLTE-EQLETHTSISAKKAKQTNGENDSQWEPQCHLVHATPYESRDIAVLQN 247
 DB 99 -KXALVQIYELEEHKTIETWREYVQ-----DSFKP---LVCISNAPLDFVAVSLRN 148

QY 248 GISTPVIYSSSDSGFPQLLHSLGILKICRYFQNSGTNLPLNQPCVSIPIGSGWV 307
 DB 149 KLRPLVIDPESGN-----TLYLTHKRLKFL-KLFIFFPKPEPMKSLLELQIGTYA 202

QY 308 PKIGDINSPLMLRPNASLSSANMLVQAGVSSIFIVDNDLSLDTYSRSDITAXKK 367
 DB 203 -----NIAMVRVTTTPVVALGIFVQHRVSALEPVDKGRVVDIYSKFDVINLAAEK 253

QY 368 VYTHVELDEWTHQALQLOQDANTPGFENGQRCQCLRSDDPLLKMYERLANPGVRVFI 427
 DB 254 TNNL-----DVSVTKAQHRS-----YFEG---VLKCYLHETLETIINRLVEAEVRLVW 302

QY 428 VEAGSKRVGGIISLSLDFKLL 449
 DB 303 VDENDV-VKGIYSLSLDILQALV 323

RESULT 14

AAO18497

ID AAO18497 standard; protein; 331 AA.

XX AAO18497;

XX 11-OCT-2002 (first entry)

XX Human insulin receptor signaling modifier SEQ ID NO: 16.
 XX Human, insulin receptor signaling; insulin receptor signaling modifier;
 XX ISM; diabetes; metabolic syndrome; antidiabetic.

XX Homo sapiens.

XX WO200255664-A2.

XX 18-JUL-2002.

XX 11-JAN-2002; 2002WO-US001048.

XX 12-JAN-2001; 2001US-0261226P.

XX 12-JAN-2001; 2001US-0261303P.

XX 12-JAN-2001; 2001US-0261304P.

XX 12-JAN-2001; 2001US-0261335P.

XX 12-JAN-2001; 2001US-0261336P.

XX 12-JAN-2001; 2001US-0261361P.

XX 12-JAN-2001; 2001US-0261456P.

XX 12-JAN-2001; 2001US-0261457P.

XX 12-JAN-2001; 2001US-0261458P.

XX 12-JAN-2001; 2001US-0261459P.

XX 12-JAN-2001; 2001US-0261461P.

XX 12-JAN-2001; 2001US-0261518P.

XX 12-JAN-2001; 2001US-0261531P.

XX 12-JAN-2001; 2001US-0261532P.

XX 12-JAN-2001; 2001US-0261589P.

XX 12-JAN-2001; 2001US-0261590P.

XX 12-JAN-2001; 2001US-0261594P.

XX 12-JAN-2001; 2001US-0261595P.

XX 12-JAN-2001; 2001US-0261597P.

XX (EXEL-) EXELIXIS INC.

XX Seidel-Dugan C, Ferguson KC, Kidd T;

XX WPI; 2002-599664/64.

XX N-PSDB; AAL48616.

XX Identifying an insulin receptor signaling modulator, useful as drug
 XX targets for treating diabetes or metabolic disorders, comprises
 XX contacting an assay system comprising insulin receptor signaling
 XX modifiers with a test agent.
 XX Disclosure; Page 54-55; 232pp; English.

XX The present invention relates to a method of identifying a candidate
 XX insulin receptor (INR) signaling modulating agent, involving contacting
 XX an assay system comprising an insulin receptor signaling modifier (ISM)
 XX polypeptide or nucleic acid with a test agent, and detecting a test agent
 XX -biased activity of the assay system. The method is useful for
 XX identifying candidate INR signaling modulating agents. ISM genes may be
 XX used as drug targets for treatment of disorders related to INR signaling
 XX such as diabetes or metabolic syndrome. ISM nucleic acids and
 XX polypeptides are useful for identifying and testing agents that modulate
 XX ISM function and for other applications related to the involvement of ISM
 XX in INR signaling, and for identifying subjects having a predisposition to
 XX such diseases associated with INR signaling. The present sequence is an
 XX ISM protein described in the exemplification of the invention

XX Sequence 331 AA;

Query Match 15.7%; Score 367; DB 5; Length 331;

Best Local Similarity 29.1%; Pred. No. 4.1e-28;

Matches 111; Conservative 74; Mismatches 133; Indels 64; Gaps 14;

QY 69 LSSPSTPSRMNDVNDENFQRTVTLSDGTWSEGLTVSEAAIQISRCRVSEYLNHTCY 128
 Db 5 ISSDSSPA-----VENEHPQETPE-SNSV-----YTSFMKSHRCY 39

QY 129 DLLPDSGKVIALDINLVPKQSFHLLHEOGIPVAPLWDSFRGQFVGLLSPDLFILIRELE 189
 Db 40 DLIPTSSKLVVFDTSLVQKKAFALVTNGVRAAPLWDSKQSFVGMILTITDFINILHRY- 98

QY 189 THGSNLTE-EQLETHHTISANKEAKRQTNGRNDSQWRPQOHLVHATPYESLRDIAVKLLON 247
 Db 99 -YKSALVQIYELEHKTETWREVYLQ-----DSFKP---LVCISPNASLFDVSSILRN 148

QY 248 GISTVPVIYSSSDGSPQQLHLASLSGILKICRYFNKSTGNLPIINQPVCSIPLSWV 307
 Db 149 KIHRLPVIDPESGN-----TLVILTHKRLKFL-KLFITEFPKPEFMSKSEELQIGTYA 202

QY 308 PKIGDLNSRLPMLRPNASLSALMLVQAGVSSIPVDDNDSLLDTYSRSDITALAKDK 367
 Db 203 -----NIAMVRTTTPVYVALGIFVQHRVSALPVVDEKGRVVDIYSKFDVINLAAEK 253

QY 368 VYTHVRLDEMTIHOALQODANTPFGFNGRCQMCRLSDPLKVMERLANPGVRRVFI 427
 Db 254 TYNL---DVSVTALQHRSH-----YFEG--VLKCYLHETLETIINRLVEABVHRLVV 302

QY 428 VEAGSKRVEGIISLSDIFKFL 449
 Db 303 VDENDV-VKGIVSLSDILQALV 323

RESULT 15
 ABU04258
 ID ABU04258 standard; protein; 331 AA.
 AC ABU04258;
 XX
 DT 29-JAN-2003 (first entry)
 DE Human expressed protein tag (EPT) #924.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 FN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 30-AUG-2001; 2001US-0310601P.
 PR 31-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCO INC.
 XX
 XX Chicx RM, Tomlinson AJ, Urban RG;
 XX
 DE WPI; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX

PS Example 2; SEQ ID NO 924; 134pp; English.
 XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, or protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for transla-
 CC tional profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 331 AA;

Query Match 15.7%; Score 367; DB 6; Length 331;
 Best Local Similarity 29.1%; Pred. No. 4.1e-28;
 Matches 111; Conservative 74; Mismatches 133; Indels 64; Gaps 14;

QY 69 LSSPSTPSRMNDVNDENFQRTVTLSDGTWSEGLTVSEAAIQISRCRVSEYLNHTCY 128
 Db 5 ISSDSSPA-----VENEHPQETPE-SNSV-----YTSFMKSHRCY 39

QY 129 DLLPDSGKVIALDINLVPKQSFHLLHEOGIPVAPLWDSFRGQFVGLLSPDLFILIRELE 188
 Db 40 DLIPTSSKLVVFDTSLVQKKAFALVTNGVRAAPLWDSKQSFVGMILTITDFINILHRY- 98

QY 189 THGSNLTE-EQLETHHTISANKEAKRQTNGRNDSQWRPQOHLVHATPYESLRDIAVKLLON 247
 Db 99 -YKSALVQIYELEHKTETWREVYLQ-----DSFKP---LVCISPNASLFDVSSILRN 148

QY 248 GISTVPVIYSSSDGSPQQLHLASLSGILKICRYFNKSTGNLPIINQPVCSIPLSWV 307
 Db 149 KIHRLPVIDPESGN-----TLVILTHKRLKFL-KLFITEFPKPEFMSKSEELQIGTYA 202

QY 308 PKIGDLNSRLPMLRPNASLSALMLVQAGVSSIPVDDNDSLLDTYSRSDITALAKDK 367
 Db 203 -----NIAMVRTTTPVYVALGIFVQHRVSALPVVDEKGRVVDIYSKFDVINLAAEK 253

QY 368 VYTHVRLDEMTIHOALQODANTPFGFNGRCQMCRLSDPLKVMERLANPGVRRVFI 427
 Db 254 TYNL---DVSVTALQHRSH-----YFEG--VLKCYLHETLETIINRLVEABVHRLVV 302

QY 428 VEAGSKRVEGIISLSDIFKFL 449
 Db 303 VDENDV-VKGIVSLSDILQALV 323

Search completed: July 7, 2004, 17:50:50
 Job time : 61 secs

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	1030.5	44.0	391	2	B86222	hypothetical prote	
2	355.5	15.2	330	2	T10759	AMP-activated prot	
3	303.5	13.0	322	1	SGBYC3	regulatory protein	
4	229.5	9.8	274	2	S67444	probable 5'-AMP-ac	
5	228	9.7	478	2	T25899	hypothetical prote	
6	154	6.6	460	2	T24248	hypothetical prote	
7	123	5.3	270	2	T09514	5'-AMP-activated p	
8	122.5	5.2	443	2	T25854	hypothetical prote	
9	115	4.9	629	2	T18227	hypothetical prote	
10	114	4.9	299	1	G64453	conserved hypothet	
11	114	4.9	396	2	D96832	hypothetical prote	
12	112.5	4.8	280	1	H64452	conserved hypothet	
13	112.5	4.8	408	2	T43240	CBS-domain protein	
14	108.5	4.6	168	1	A64478	hypothetical prote	
15	108	4.6	1547	2	A12343	hypothetical prote	
16	107.5	4.6	904	2	D71603	replication factor	
17	107.5	4.6	4910	2	S64542	probable membrane	
18	107	4.6	392	2	D71175	hypothetical prote	
19	107	4.6	476	2	AB1294	glutamyl-tRNA (Gln)	
20	106.5	4.5	567	2	D72359	conserved hypothet	
21	106	4.5	703	2	AH1665	glutamyl-tRNA (Gln)	
22	106	4.5	703	2	T43557	F-box/WD-repeat pr	
23	106	4.5	842	2	C83177	probable phosphor	
24	106	4.5	1475	2	T33318	hypothetical prote	
25	105	4.5	527	2	S64060	probable membrane	
26	104.5	4.5	514	1	B69214	MDJ0100 protein hom	
27	104.5	4.5	781	2	B64135	ribonucleoside-dip	
28	103.5	4.4	989	1	D69030	MJ1225-related pro	
29	103.5	4.4	261	2	T02842	probable membrane	

Gene 67, 247-257, 1988
A>Title: Molecular characterization of yeast regulatory gene CAT3 necessary for glucocorticoid-induced expression of the *hprt* gene
A:Reference number: J70316; MUID:89006284; PMID:3049255
A:Accession: J70316
A:Molecule type: DNA
A:Residues: 1-322 <SCH>
A:Cross-references: GB:M21760; NID:gl71164; PIDN:AAA34472.1; PID:gl71165
R:DOI, A.; DOI, K.
submitted to the EMBL Data Library, June 1993
A:Description: Correct end of the ORF for the CDC20 gene of *Saccharomyces cerevisiae*.
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T10759
R:Woods, A.; Cheung P. C.F.; Smith, F.C.; Davison, M.D.; Scott, J.; Beri, R.K.; Carling, J. Biol. Chem. 271, 13282-13290, 1996
A>Title: Characterization of AMP-activated protein kinase beta and gamma subunits: Assembly and function
A:Reference number: Z36738; MUID:96215327; PMID:8626596
A:Accession: T10759
A:Status: translated from GE/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-330 <WCG>
A:Cross-references: EMBL:X95578; NID:gl1185270; PIDN:CAA64831.1; PID:gl1185271
A:Experimental source: strain Wistar
C:Complex: heterotrimer; alpha, beta and gamma chains
C:Function:
A:Description: is responsible for the regulation of fatty acid synthesis by phosphorylation of acetyl-CoA carboxylase
C:Superfamily: CAT3 protein
C:Keywords: fatty acid biosynthesis; phosphotransferase

Query Match 15.2%; Score 355.5; DB 2; Length 330;
Best Local Similarity 30.3%; Pred. No. 6.5e-20;
Matches 100; Conservative 67; Mismatches 124; Indels 39; Gaps 11;
Query 121 YLMHCTCYLLDPGSKVIALDINLPVKQSHILHEOGIPVAPLWDSFGCFVGLLSPLDF 180
Db 31 FMKSHRCYDLIPTSSKLVDFVTSQVKAFFALVINGVRAAPLWDSKQSFVGMLTITDF 90
Query 181 ILILRELETHGSLNTE-EGLETHHTISAWKEAKRQTNGRNDSQWRPQOHLVHATPVESLRD 239
Db 91 INILHRY--YKSAALVQVLEEHKHTWREVLQ-----DSFKP-----LVCIISPASLFD 139
Query 240 IAVKLLONGISTVPVIYSSSDGSPQLLHLASLSGLKACICRYEKNSTGNLPILNPQVC 299
Db 140 AVSLIRNKHRLPVIDPDSGN-----TUIYTHKRLKFL-KLFTPEPKPEFMSKLE 193
Query 300 SIPLGSWPKIGDLSNRPLAMLRPNASLSALNMLVQAGVSSIPVDNDSELDITSRSD 359
Db 194 ELQIGTYA-----NIAMVKTTPVVALGIFVQHRVSALPVVDKRGVVDIYSKFD 244
Query 360 ITALAKDKVYTHVRLDEMTHIHOALQIGQDANTPPFFGNGQRCQMLRSDPLIKWLERLAN 419
Db 245 VINAAEKTYNNL--DVSVTALQRRSH-----YFEG--VLKCYLHETLEAIIINRLVE 293
Query 420 PGVRRVFIVEAGSKRVEGIISLSDIFKFL 449
Db 294 AEVHLVWVDEHDV-VKGIIVSLSDILQALV 322

RESULT 3
RBYC3
regulatory protein SNF4 - Yeast (*Saccharomyces cerevisiae*)
N/Alternate names: CAT3 protein; protein G2945; protein YGL115W
C:Species: *Saccharomyces cerevisiae*
C:Date: 30-Jun-1989 #sequence_revision 10-Jun-1989 #text_change 21-Jul-2000
C:Accession: A38906; J70316; S48508; S64125
R:Celenza, J.B.; Eng, F.J.; Carlson, M.
Mol. Cell. Biol. 9, 5045-5054, 1989
A>Title: Molecular analysis of the SNF4 gene of *Saccharomyces cerevisiae*: evidence for a role in the regulation of the *hprt* gene
A:Reference number: A33480; MUID:90097921; PMID:2481228
A:Accession: A38906
A:Molecule type: DNA
A:Residues: 1-322 <CEL>
A:Cross-references: GB:M30470; NID:gl172635; PIDN:AAA35061.1; PID:gl172636
R:Schueller, H.G.; Entian, K.D.

Gene 67, 247-257, 1988
A>Title: Molecular characterization of yeast regulatory gene CAT3 necessary for glucocorticoid-induced expression of the *hprt* gene
A:Reference number: J70316; MUID:89006284; PMID:3049255
A:Accession: J70316
A:Molecule type: DNA
A:Residues: 1-322 <SCH>
A:Cross-references: GB:M21760; NID:gl71164; PIDN:AAA34472.1; PID:gl71165
R:DOI, A.; DOI, K.
submitted to the EMBL Data Library, June 1993
A:Description: Correct end of the ORF for the CDC20 gene of *Saccharomyces cerevisiae*.
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T10759
R:Woods, A.; Cheung P. C.F.; Smith, F.C.; Davison, M.D.; Scott, J.; Beri, R.K.; Carling, J. Biol. Chem. 271, 13282-13290, 1996
A>Title: Characterization of AMP-activated protein kinase beta and gamma subunits: Assembly and function
A:Reference number: Z36738; MUID:96215327; PMID:8626596
A:Accession: T10759
A:Status: translated from GE/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-330 <WCG>
A:Cross-references: EMBL:X95578; NID:gl1185270; PIDN:CAA64831.1; PID:gl1185271
A:Experimental source: strain Wistar
C:Complex: heterotrimer; alpha, beta and gamma chains
C:Function:
A:Description: is responsible for the regulation of fatty acid synthesis by phosphorylation of acetyl-CoA carboxylase
C:Superfamily: CAT3 protein
C:Keywords: fatty acid biosynthesis; phosphotransferase

Query Match 13.0%; Score 303.5; DB 1; Length 322;
Best Local Similarity 27.9%; Pred. No. 6.8e-16;
Matches 95; Conservative 64; Mismatches 129; Indels 53; Gaps 12;
Query 118 VSEVINLHCTCYLLDPGSKVIALDINLPVKQSHILHEOGIPVAPLWDSFGCFVGLLSPL 177
Db 22 IRKFLNSKTSDYDLPVPSYRLVLDTSLLVKKSLNVLQNSIVSAPLWDSKTSREAGLLTT 81
Query 178 LDFTILRLLETHGSLNTE-EGLETHHTISAWKEAKRQTNGRNDSQWR---PQOHLVHATPY 234
Db 82 TDFINV---IQYFSPNDKFLVD-----KLQDGLKDIERALGVDDLTASHP 129
Query 235 ESLEADIAVKLLONGISTVPVIYSSSDGSPQLLHLASLSGLKCI---CR---YFKNST 288
Db 130 RPLFEACLKMLESRSGRIPLI-DQDETHREIVSVSVLTQVRLKFLVALNCRHFLKIP 188
Query 289 GNLPILNPQVCSTPLGSWPKIGDLSNRPLAMLRPNASLSALNMLVQAGVSSIPVDND 348
Db 189 GDLNIIQD-----ANWSCQMTTP---VIDVIMQITQGRVSVSVPIDEN 229
Query 349 DSLDITYSRSDITALAKDKVYTHVRLDEMTHIHOALQIGQDANTPPFFGNGQRCQMLRSD 408
Db 230 GYLINVEAYDVLGLIKGGIYNDLSL---SVGEALMRRSD-----FEG--VYTCTKND 278
Query 409 PLLKWMERLANPGVRRVFIVEAGSKRVEGIISLSDIFKFL 449
Db 279 KLSTIMDNIRKARVHRFFVVD-DVGRVLGVLTSLDILKYL 318

RESULT 4
S67444
probable 5' AMP-activated, gamma subunit family - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: T38059; S67444
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z21766
A:Accession: T38059
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-274 <MC2>
A:Cross-references: EMBL:Z69944; NID:G1217974; PID:CAA93805.1; PID:G1217975; GSPDB:GN000200
C:Genetics:
A:Map position: 1
A:Introns: 12/3; 54/1; 173/3; 226/1
C:Superfamily: CAT3 protein

Query Match 9.8%; Score 229.5; DB 2; Length 274;
Best Local Similarity 25.4%; Pred. No. 2.9e-10;
Matches 72; Conservative 60; Mismatches 118; Indels 33; Gaps 10;

QY 118 VSEYLNLFHPCYDLPDPSGKVIADNLNPKQSFHILHEQGIPIVAPLWDSFGQFVGLLSP 177
DB 15 IQAFIRSRYSYDVLTFSEFLVDFVTLFVKTSLSLLTLNNIVSAPLWDSSEANKFACILLTM 74

QY 178 LDFILILRELETHGSLNLT-EOLETHHTISAWKEAKQTNGRNDSCWRPQQHLVHATPYE 235
DB 75 ADFVNVIK-YYQSSSFPFAIAEIDKFRGLGUREVERKIGA-----IPETIYVH--PMH 126

QY 236 SLRDIAVXLLQNGISTVPVYSSSDGSPFLLHSLASGLIKKICRYFKNSTGNLPILN 295
DB 127 SLMDACLAMSKSRARRIPLIDVDGETGS-EMIVSVLTQYRIKF-SMNCKETA---MLR 181

QY 296 QVCSIPDGSWPKIGDNLNRPPLNPNASLSSALNMLVQAGVSSIPVDDNDSDLLDTY 355
DB 182 VPLNQMTIGTW-----SNLATASMETKVDYVIXMLAERNISGAVPVNSEGTLLNVY 232

QY 356 SRSDITALAKQKVYTHVRLEDEMTHQALQIGODANTPFGEFNG 398
DB 233 ESDVWHLIQDGYNSL---DUSVGEAL-LKRPAN-----FDG 266

RESULT 5
T25899
hypothetical protein T20F7.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T25899
R:Miller, N.; Gattung, S.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid T20F7.
A:Reference number: Z20107
A:Accession: T25899
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-478 <MIL>
A:Cross-references: EMBL:U97550; PID:AA852856.1; GSPDB:GN00028; CESP:T20F7.6
A:Experimental source: strain Bristol N2; clone T20F7
C:Genetics:
A:Gene: CESP:T20F7.6
A:Map position: X
A:Introns: 33/3; 112/2; 144/1; 205/3; 263/2; 300/3; 380/3; 402/2; 451/1

Query Match 9.7%; Score 228; DB 2; Length 478;
Best Local Similarity 25.1%; Pred. No. 8.8e-10;
Matches 94; Conservative 74; Mismatches 148; Indels 58; Gaps 16;

QY 91 TVTLSDGTVSEGLRVSEAAIISRCRVSEYLNLTGYLLPDSGKVIADNLNPKQSF 150
DB 22 TWTESDEVLPK-TPNDKEAFSL-----LWVNCQYEAMPSSSNQVDFDQGLLMHKAF 72

QY 151 HILHEQGIPIVAPLWD-SFGQFVGLLSPDLFILIREL-----ETHGSLNLTTE 197
DB 73 NGLLAQSTHVLNLLSDPFDGKGLDGLSVTFDKVWLKIVRETKCKESTELDMTQIANE 132

QY 198 QLETHHTISAWKEAKQTNGRNDSCWRPQQHLVHATPYESLRDIANKLLONGISTVPVYS 257
DB 133 EIGNUSISQYRELVRK-----EGNURP---LVSVDASSLLDAACILAEHRVHRIPVI-- 182

QY 258 SSSDGSFPQLHLASLGLIKKICRYFKNSTGNLPINQVCSIPDGSWPKIGDNLNRP 317
DB 183 DFLDGS---AFILTHKRILKFLWLFQKH-LAPLEYLKHSPKELGIGTW-----SG 229

Query Match 6.6%; Score 154; DB 2; Length 460;
Best Local Similarity 18.9%; Pred. No. 0.00046;
Matches 63; Conservative 81; Mismatches 134; Indels 56; Gaps 10;

QY 120 EYLNLTCHTCDLLPDSGKVIADNLNPKQSFHILHEQGIPIVAPLWDSFGQFV--GLLSP 177
DB 141 QYMSVVDVVELCENNSKVIIDASTTTAFRIMRDHNTTLIVWTSDBRHKRNILTL 200

QY 178 LDFILILRELETHGSLNLTBEQLETHHTISAWKEAKQTNGRNDSCWRPQQHLVHATPYE 237
DB 201 TDCLNALR-----NETPPADGQVLRASDILSGNQ-----LVSVSISSKI 239

QY 238 RDIANKLLONGISTVPVYSSSDGSPFQLHLASLGLIKKICRYFKNSTGNLPILNQP 297
DB 240 LDCHECHQNRHLRVVVLDDAK-----EVNIIISVRRVIAALHKQ-NRSLHFAQLMSKS 292

QY 298 VCSIPDGSWPKIGDNLNRPPLNPNASLSSALNMLVQAGVSSIPVDDNDSDLLDTYSR 357
DB 293 IGMSAIGTW-----ENVAVISQNETVYRAMEDMLGHYSALPVDVSKQNVIGVITK 343

QY 358 SDI-TALAKQKVYTHVRLEDEMTHQALQIGODANTPFGEFNGQRCOMCLRS-DPLAKVME 415
DB 344 TDICAPLNPFEPKRWLOETKVSIDLHT-----CKSQILLISADSAGQVLD 390

QY 416 RLANKPVRVFFIVEAGSKRVEGLISLSDIFPKLL 449
DB 391 TLLAGTQSAFPAHNG--KAIGVISLTDFLSHIL 422

RESULT 7
T09514
5'-AMP-activated protein kinase (EC 2.7.1.1-) beta-1 chain - human
C:Species: Homo sapiens (man)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T09514
R:Stapleton, D.; Woollatt, E.; Mitchell, K.; Nicholl, J.K.; Fernandez, C.S.; Michel;
FEBS Lett. 409, 452-456, 1997
A:Title: AMP-activated protein kinase isoenzyme family: subunit structure and chromoso
A:Reference number: Z16707; MUID:97367941; PMID:9224708
A:Accession: T39514

Db 361 RVHRVWVYDQNG-GLQGLVSLTDIIAVWRS 389

RESULT 12

H64452

conserved hypothetical protein MJ1225 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: H64452

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.M.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, A.; Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999; PMID:8668087

A:Accession: H64452

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-280 <BUL>

A:Cross-references: GB:U67563; GB:L77117; NID:g2826379; PIDN:AAB99228.1; PID:gl591856;

C:Genetics:

A:Map position: FOR1167099-1167941

C:Superfamily: conserved hypothetical protein MJ1225; CBS homology

F:157-204/Domain: CBS homology <CBS>

Query Match 4.8%; Score 112.5; DB 1; Length 280;

Best Local Similarity 20.6%; Pred.No.0.37;

Matches 46; Conservative 41; Mismatches 101; Indels 35; Gaps 6;

Qy 228 LVHATPPYSLRDIAVKLLQNGISTVPVIYSSSD--GSPQLHLHLASLGLKLCICRYFK 285

Db 14 IVTVPTTIRKALMTNENKVRLLPVNAGNKVVGIIITSDIVDFMGSSKYNLIREK 73

Qy 286 NSTGNLPILNPWCSPICGSWVPKTDGNSRPLMRPNASLSALNMLIVQAGVSSIPIV 345

Db 74 HERNFLAAINFV-----REIMEENVITIKENADIDEAETETITKNVGGAPIV 121

Qy 346 DDNLSLIDTYSRSDITALAKDKVYTHVSLDEMTIHQALQLGQDANTPFGFNGRCQMCL 405

Db 122 NDEQLSLITERVIRALLKDENEVIDDVIIRDV-----VATP-----GER----- 166

Qy 406 RSDPLLKVMERLANPGVRVFTVEAGSKRVEGIIISLSDIPKFL 448

Db 167 ---LKDVRTVMVRNGFRLLPVWSEG--RLVGIIITSTDFIKLL 203

RESULT 13

T43240

CBS-domain protein Sds23p - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T43240; T40589; T43242

R:Ishii, K.; Kumada, K.; Toda, T.; Yanagida, M. EMBO J. 15, 6629-6640, 1996

A:Title: Requirement of Ppl phosphatase and 20S cyclosome/APC for the onset of anaphase

A:Reference number: Z22359; MUID:97133292; PMID:8978689

A:Accession: T43240

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-408 <ISH>

A:Cross-references: EMBL:D86840; NID:gl486256; PIDN:BAAL13172.1; PID:gl486257

R:Seeger, K.; Harris, B.L.; Wood, V.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, January 1999

A:Reference number: Z21938

A:Accession: T40589

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-408 <SEE>

A:Cross-references: EMBL:AL035216; PIDN:CAA22817.1; GSPDB:GN00067; SPDB:SPBC646.13

A:Experimental source: strain 972h-; cosmid c646

R:Kawamukai, M.

Biochim. Biophys. Acta 1446, 93-101, 1999
 A>Title: Isolation of a novel gene, moc2, encoding a putative RNA helicase as a suppressor of the growth defect of a yeast mutant.
 A:Reference number: Z22361; PMID:99326140; PMID:10395922
 A:Accession: T43242
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-408 <RAW>
 A:Cross-references: EMBL:D87870; NID:gl565198; PIDN:BAAL3486.1; PID:gl565199
 A:Gene: SPBC646.13
 A:Map position: 2
 A:Note: sds23

Query Match 4.8%; Score 112.5; DB 2; Length 408;
 Best Local Similarity 20.6%; Pred. No. 0.64;
 Matches 75; Conservative 57; Mismatches 121; Indels 111; Gaps 16;
 QY 156 QGTPVAPLWDSFRGQFVGLLSP-----LQFILLIRELE-----THGSN----- 193
 DB 43 QDIPVTFLDHN-----ETALIDPETSMEERASSILLDRLSALPIVAAGSNEIATTFPYAD 98
 QY 194 -----LATEQLETHITSAWKEAKRQTRGRNDSONRPOQHLVHAT 232
 DB 99 LNSFLMWVGFDDFNDGRFKKVAEDIRAGKVITAVEVAKL--GKNKDDFIT--IPHTT 152
 QY 233 PYESLSDIAVKLLONGISTVPVIYSSSDGSPQOLLHLASLSGLIKICICRYKNSGTLNLP 292
 DB 153 SLGLRLAEI-----LSSGIRRVAV-----TNEQG-----ELSFMSQRSII-----RFLNNIRAFP 198
 QY 293 ILNQPVCSIFLGSWVPKIGDLSRPLAMLRPNASLSALLNMLVQAGVSSIPVINDNSLL 352
 DB 199 DL-EPLMSRTIHSI-----DIGSTDITCISGQKVAALRQMQVTGIGSLAVVDAQFRL 252
 QY 353 DTYSRSDITALAK-----DKVYTHVRLDMTHQALQIQDANTPFPGFNGQRCOMCL 405
 DB 253 GNISLVQVKKVTRSSSVYLLNKSCHAF--LSVKSQEGIRAGKLSAPAFNIY----- 302
 QY 406 RSDPLKWMERLANPGRVRFVFEAGS-----KRVGGIISLSDI 444
 DB 303 ESSTFAFTAKIVATQCHRLWLVQSPSCPPSPKNAHLSPGSMGVKVNQLLVGVSLTDI 362
 QY 445 FKFL 448
 DB 363 ISVL 366

RESULT 14
 A64478
 hypothetical protein homolog MJ1426 - Methanococcus jannaschii
 C:Species: Methanococcus jannaschii
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: A64478
 R:Bult, C.J.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Bult, C.J.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Rison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kake, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; PMID:96337999; PMID:8688087
 A:Accession: A64478
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-168 <BL>
 A:Cross-references: GB:U67583; GB:L77117; NID:g2826409; PIDN:AAB99437.1; PID:gl592076; T5;120-168/Domain: CBS homology <CBS>
 C:Genetics:
 A:Map position: FOR1396048-1396554
 A:Start codon: TTG
 C:Superfamily: conserved hypothetical protein yhcV; CBS homology
 F;120-168/Domain: CBS homology <CBS>

Query Match 4.6%; Score 108.5; DB 1; Length 168;
 Best Local Similarity 26.2%; Pred. No. 0.35;
 Matches 38; Conservative 32; Mismatches 48; Indels 27; Gaps 7;

QY 234 YE--SLRDIAVKLLONGISTVPVIYSSSDGSPQOLLHLASLSGLIKICICRYKNSGTLN 291
 DB 28 YEDNDLIDVIRLFRKNKISGAPVL--NKDG--KLWGISSEDIIVTIIVTH--NEDLNL 79
 QY 292 PILNQP--VCSIPL-----GSWVPKIGDLSRPLAMLRPNASLSALLNMLV 335
 DB 80 -ILPSPDLIELPLKALKIEEPMEDLKNALKTKVRDVMTRKVIKIVAPDMTINDAAKLWV 138
 QY 336 QAGVSSIPVDDNDNSLLDTYSRSDI 360
 DB 139 KNNIKRLPVVDDEGNLIGIVTRGDL 163

RESULT 15
 A12843
 hypothetical protein alr193 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change C9-Dec-2002
 C:Accession: A12843
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigun, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
 A:Reference number: AB1807; PMID:21595285; PMID:11759840
 A:Accession: A12843
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1547 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAH73602.1; PID:gl7130993; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr193

Query Match 4.6%; Score 108; DB 2; Length 1547;
 Best Local Similarity 21.0%; Pred. No. 10;
 Matches 81; Conservative 62; Mismatches 173; Indels 70; Gaps 15;
 QY 98 TVSEGTURVSEALQISRCV-----SEYLNHTCYDLLPSGKVIALDINKLPVQSPHIL 153
 DB 48 TVQSAKEERSASQEVRLDVLAKRKYAOEHVLLVGRPGSGKSTAL--LRLLEAEKRL 106
 QY 154 -----HEQIPVAPLWDSFRGQFVGLLSPDLFTLILRELETHGSNLTETEQTHTISAW 207
 DB 107 POPLDAERGVLAPP---SLAGKGVGGIQLPILVELRYVYQTSILDLIRDFLKHCLLD 163
 QY 208 KEAKRQTRNGRNDSONRPOQHLVHATPYESLRDIAVKLLONGISTVPVIYSS-----SSDG 262
 DB 164 TATIEQL--LFDGQFLLVDGINELPSEKARQDLKFERQDNQKTPMIFTTRDLGVGGDL 221
 QY 263 SFPOLLHLASLSG--ILKICRYEK-----NSTGNLPILNQPVCS--SIFLG 304
 DB 222 GITKLEMQPLTAEQVQFVRAVLPAGEOMLQNLGRLREFGETPILLFMLCLLFFVETG 281
 QY 305 SWVPKIG--DLNSRPLAMLRPNASLS-----ALNMLVQAGVSSIPVDD 347
 DB 282 DIPNLGLLFRELQRYKNELKPNVPKGVSRDHSWMLQLAFNM-----TKGDKLTEL 336
 QY 348 NDSLDTYSRSDITALAKQVYTHV-----LDENIHOALQIQDANTPFPGFNGQRC 401
 DB 337 NVAILKSKAEVITQLLKDEGNFPRDQASELSKDLNHLIQLA--ANDQIE-FRHQLI 393
 QY 402 QMCLRSPLKWMERLANPGRVRFI 427
 DB 394 QEYTYACLLKLLPSLSDSLQREYL 419

Search completed: July 7, 2004, 17:52:51
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:47:08 ; Search time 17 Seconds
(without alignments)

1381.391 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREHLPMSPIEGPTVFQAI.....SKRVEGISLSDFKELLSL 451

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	373	15.9	569	1 AAKH HUMAN	Q9ugj0 homo sapien
2	367	15.7	331	1 AAKG HUMAN	P54659 homo sapien
3	365	15.6	330	1 AAKG BOVIN	P58108 bos taurus
4	355.5	15.2	330	1 AAKG RAT	P80385 rattus norv
5	350	14.9	330	1 AAKG MOUSE	O54950 mus musculus
6	340	14.5	464	1 AAKI HUMAN	Q9ugi9 homo sapien
7	334.5	14.3	514	1 AAKI PIG	Q9mfp4 sus scrofa
8	303.5	13.0	322	1 SNF4 YEAST	P12904 saccharomyc
9	289	12.3	334	1 YL28 SCHPO	Q10343 schizosacch
10	261.5	11.2	328	1 SNF4 KLULA	Q9p859 kluyveromyc
11	124	5.3	269	1 AAKB MOUSE	Q9r078 mus musculus
12	124	5.3	269	1 AAKB RAT	P80386 rattus norv
13	123	5.3	269	1 AAKB HUMAN	Q9y478 homo sapien
14	122.5	5.2	271	1 AAKB RAT	Q9qzh4 rattus norv
15	121	5.2	272	1 AAKC HUMAN	Q43741 homo sapien
16	116	5.0	122	1 AAKB PIG	P80387 sus scrofa
17	114	4.9	296	1 YC32 METJA	Q58629 methanococc
18	112.5	4.8	280	1 YC25 METJA	O58622 methanococc
19	112.5	4.8	408	1 SD23 SCHPO	O09826 schizosacch
20	108.5	4.6	168	1 YE26 METJA	Q58821 methanococc
21	107.5	4.6	4913	1 MDN1 YEAST	Q12019 saccharomyc
22	107	4.6	476	1 GATE LISMO	Q8y6d3 listeria mo
23	106.5	4.5	548	1 YPAC THEMA	Q9wz56 thermotoga
24	106	4.5	476	1 GATB LISIN	Q92aq4 listeria in
25	106	4.5	703	1 POP2 SCHPO	O14170 schizosacch
26	105	4.5	527	1 YGF6 YEAST	P43172 saccharomyc
27	104.5	4.5	756	1 RIRI HAEIN	P43754 haemophilus
28	101.5	4.3	302	1 YR33 THEPE	P15889 thermofilum
29	101	4.3	509	1 Y100 METJA	Q57564 methanococc
30	100.5	4.3	958	1 MSH4 MOUSE	O99mt2 mus musculus
31	99.5	4.2	1024	1 SRE1 RAT	P56720 rattus norv
32	98.5	4.2	3343	1 YOG7 CABEL	P34616 caenorhabdi
33	97.5	4.2	238	1 YD43 HAEIN	P71379 haemophilus

RESULT 1

ID	AAKH HUMAN	STANDARD	PRT	569 AA
AC	Q9UGJ0: Q9UGJ0; Q9UDN8; Q9ULX8;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	5'-AMP-activated protein kinase, gamma-2 subunit (AMPK gamma-2 chain)			
DE	(AMPK gamma2) (H91620p).			
GN	PRKAG2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RX	MEDLINE=20164049; PubMed=10698692;			
RA	Cheung P.C.F., Salt I.P., Davies S.P., Hardie J.G., Carling D.;			
RT	Characterization of AMP-activated protein kinase gamma-subunit			
RT	isoforms and their role in AMP binding."			
RL	Biochem. J. 346:659-669(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM B).			
RX	MEDLINE=20564210; PubMed=1112354;			
RA	Lang T., Yu L., Qiang T., Jiang J., Chen Z., Xin Y., Liu G., Zhao S.;			
RT	Molecular cloning, genomic organization, and mapping of PRKAG2, a			
RT	heart abundant gamma-2 subunit of 5'-AMP-activated protein kinase, to			
RL	human chromosome 7q36."			
RL	Genomics 70:258-263(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM B).			
RC	TISSUE=Placenta;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Orsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,			
RA	Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,			
RA	Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;			
RT	"NEDO human cDNA sequencing project."			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM B).			
RC	TISSUE=Liver;			
RA	MEDLINE=22398257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullay S.C.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hlyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			

34	97.5	4.2	701	1 YL66 YEAST
35	97.5	4.2	2052	1 FYV1 MOUSE
36	97	4.1	196	1 Y525 METKA
37	97	4.1	909	1 NUOG SHEON
38	96	4.1	551	1 IL2B HUMAN
39	96	4.1	2037	1 FASI CANAL
40	95.5	4.1	656	1 MUTL LACIA
41	95.5	4.1	977	1 YAG9 SCHPO
42	95	4.1	690	1 NCPR PHAHO
43	94	4.0	465	1 YGT4 YEAST
44	94	4.0	527	1 YB64 YEAST
45	94	4.0	951	1 LGR4 RAT

ALIGNMENTS


```
Db 246 LFEDEAVESSESGVYMRFRSHKCYDIVPTSSKLVFDTTLQVKKAFVALVANGVRAAP 305
QY 163 LNDSPRGQVGLSPDLFILRELETHGNSLITE-BOLETHHTISAMKEAKRQTNGRNDSQ 221
Db 306 LNESKQSFVGMUTITDFINLHRY--YKSPWQVYELEHKEHKTETWRELYLQ-----ET 357
QY 222 WRPQOHLVHATPVESLURDIJAVKLQNGISTVPVIYSSSDSGSPQLHLASLSGLKLCIC 281
Db 358 KXP---LVNISPDASLFDVAVSLIKKHRLPVIDPISGNA-----LYILTKHRLKFL- 408
QY 282 RYFQNSTGNLPILNQVPCSPILGSPWPKIGDLNRLPLMRPNASLSSALNMLVQGVSS 341
Db 409 QLFMSDMPKPAFMKQNLDELIGTY-----HNIAFTHTPTPIKALNIFVERRISA 459
QY 342 IPTVDNDSLLTYSRSDITALAKQVYTHVRLEDMTHQALQQLGDAN--PFGFENGQRC 401
Db 460 LPVVDSESGKVVDIYSFEDVINLAETYNL-----DITVTOALQHRSQ-----YFEG--V 508
QY 402 QMCIRSDPILKVMERLANPQVRRVTV-EAGSKRVEGIISLSDFIKFL 449
Db 509 VXCNKLEILITVIDRVRAEVRHLVVVNEADS--IVGIISLSLQAL 555
[1]
RESULT 2
ID AAKG HUMAN STANDARD; PRT; 331 AA.
AC P54619;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPKg).
GN PRKAG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Fetal liver;
RX MEDLINE=96224074; PubMed=8621499;
RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,
RA Dyck J.R.B., Kemp B.E., Witters L.A.;
RT "Non-catalytic beta- and gamma-subunit isoforms of the 5'-AMP-
RT activated protein kinase."
RL J. Biol. Chem. 271:8675-8681(1996).
RN [2]
SEQUENCE FROM N.A.
RP TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Caranci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Scherch A., Schein J.E., Jones S.D.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
```

```
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U42412; AAC50495.1; -.
DR EMBL; BC000358; AAH00358.1; -.
DR Genew; HGNC:9385; PRKAG1.
DR MIM; 602742; -.
DR GO; GO:0004691; P:AMP-dependent protein kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
DR InterPro; IPR000644; CBS_domain.
DR SMART; SM00116; CBS; 4.
DR Ffam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 48 96 CBS 1.
FT DOMAIN 123 177 CBS 2.
FT DOMAIN 197 250 CBS 3.
FT DOMAIN 271 323 CBS 4.
SQ SEQUENCE 331 AA; 37579 MW; 0F22B9CA1DBD87AE CRC64;
Query Match 15.7%; Score 367; DB 1; Length 331;
Best Local Similarity 29.1%; Pred. No. 2.5e-20;
Matches 111; Conservative 74; Mismatches 133; Indels 64; Gaps 14;
QY 69 LSSPTSPGSRMNMVDNENFQRTVTLSDGTVSEGLVASEAAIQRVSEYLNLTQY 128
Db 5 ISSDSPA-----VNEHQETPE-SNSV-----YTSFMSHCY 39
QY 129 DLLPDSGKVIADINLPVKOSFHILHEQGVIPAPLWDSFRGQVGLLSPDLFILRELE 188
Db 40 DLIPTSSKLVFDTSLQVKKAFVALVANGVRAAPLWDSKQSFVGMUTITDFINLHRY- 98
QY 189 THGNSLITE-BOLETHHTISAMKEAKRQTNGRNDSQWRPQOHLVHATPVESLURDIJAVKLQN 247
Db 99 -YKSPWQVYELEHKEHKTETWRELYLQ-----DSFKP---LWCISPNASLFDVAVSSLRN 148
QY 248 GISTVPVIYSSSDSGSPQLHLASLSGLKLCICRYFKNSTGNLPILNQVPCSPILGSWV 307
Db 149 KIHRLPVIDPESGN-----LYILTKHRLKFL-KLFITEFPKPEFMSKSLLEELQIGTYA 202
QY 308 PKIGDLNRLPLMRPNASLSSALNMLVQGVSSIPVDNDSLLDTYSRSDITALAKDK 367
Db 203 -----NIAMVRITTPVYVALGIFVCHRVASLPVDEKGRVWDIYKSFVDINLAEEK 253
QY 368 VYTHVRLEDMTHQALQQLGDANTPFGFENGQRCQCLRSDDPILKVMERLANPQVRRVFI 427
Db 254 TYNL---DVSVTKALQHRSH-----YFEG---VLKCYLHETLTETINLRVRAEVRHLVV 302
QY 428 VRAGSKRVEGIISLSDFIKFL 449
Db 303 VDENVV-KGIVSLSDILQALV 323
RESULT 3
AAKG BOVIN STANDARD; PRT; 330 AA.
ID AAKG BOVIN
AC P58108;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
```

10-OCT-2003 (Rel. 42, Last annotation update)
5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
(AMPKg).
GN PRKAG1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Benkel B., Kollers S., Fries R., Szaczov A., Yoshida E., Davoren J.,
RA Hickey D.,
RT "Characterization of the bovine AMPK gamma-1 gene."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL; AF329081; AAK19307.1; -
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 48 96 CBS 1.
FT DOMAIN 123 177 CBS 2.
FT DOMAIN 197 250 CBS 3.
FT DOMAIN 271 323 CBS 4.
SQ SEQUENCE 330 AA; 37481 MW; F130ACBEE2BFE89 CRC64;
Query Match 15.6%; Score 365; DB 1; Length 330;
Best Local Similarity 28.8%; Pred. No. 3.5e-20;
Matches 109; Conservative 73; Mismatches 138; Indels 58; Gaps 13;
QY 73 STPGSRMMDVDNEFQFTVTLSDCTVSEGLRVSEMAIQISRCRVSEVNLHTCYDILP 132
DB 3 AVPSDSYDPAVENEHLQETPE-SNSV-----YTSFMKSHRCYDILP 43
QY 133 DSGKVIALDINLPVQKSHFILHEQIPVAPLWDSFGQFVGLLSPLDILILRELETHGS 192
DB 44 TSSKLAVFTSLQVKKAFALVTNGVRAAPLWDSKKQSPVGMLTITDFINILHRY--YKS 101
QY 193 NLTE-EQLTHTISAWKEAKQTNGRNDSSQWRPQCHLVHATPYESLRDIANVLLONGIST 251
DB 102 ALVQYIELEHKEIETWREVFLO-----DSFKP---LVCISPNASLFDVAVSLIRNKIHR 152
QY 252 VPVIYSSSDSGSFPQLLHSLASLGILKICICRYFKNSTGNLPLNCPVCSIPGSGVWPVKIG 311
DB 153 LPVIDPESGN-----TVILTHKRLKFL-KLFTTEFPKPEPMKSLBELQIGTYA---- 202
QY 312 DLNSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNPSLLDTYGRSDITALAKDKVYTH 371
DB 203 -----NIAMVRTTTPYVALGIFVQHRVSLPVPVDEKGRVDIYKSFVINLAEKTYNN 257
QY 372 VLDEMTHQALQLQGDQANTPEGFNGQRCQCLSRDPLLKMERLANGPVTRVIVZAG 431
DB 256 L---DVSVTKALQHRSH-----YFEG---VLKCYLHETLETIINRLVEAEVRLVVVDEN 306

QY 432 SKRVEGIISLSDIFKFL 443
DB 307 DV-VKGIIVSLSDILQALV 323
RESULT 4
ID AAKG RAT STANDARD; PRT; 330 AA.
AC P80385;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPKg).
GN PRKAG1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Wistar;
RX MEDLINE=96215327; PubMed=8626596;
RA Woods A., Cheung P.C.F., Smith F.C., Davison M.D., Scott J.,
RA Beri R.K., Carling D.,
RT "Characterization of AMP-activated protein kinase beta and gamma
subunits. Assembly of the heterotrimeric complex in vitro."
RL J. Biol. Chem. 271:10282-10290(1996).
RN [2]
RP SEQUENCE OF 8-330 FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96224074; PubMed=8621499;
RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,
RA Dyck J.R.B., Kemp B.E., Witters L.A.;
RT "Non-catalytic beta- and gamma-subunit isoforms of the
5'-AMP-activated protein kinase."
RL J. Biol. Chem. 271:8675-8681(1996).
RN [3]
RP SEQUENCE OF 48-330 FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95050763; PubMed=7961907;
RA Stapleton D., Gao G., Michell B.J., Widmer J., Mitchell K.,
RA Teh T., House C.M., Witters L.A., Kemp B.E.;
RT "Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are
homologs of proteins that interact with yeast Snf1 protein kinase."
RL J. Biol. Chem. 269:29343-29346(1994).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HORMONE-SENSITIVE LIPASE AND
CC HYDROXYMETHYLGUTARYL-COA REDUCTASE. APPEARS TO ACT AS A METABOLIC
CC STRESS-SENSING PROTEIN KINASE SWITCHING OFF BIOSYNTHETIC PATHWAYS
CC WHEN CELLULAR ATP LEVELS ARE DEPLETED AND WHEN 5'-AMP RISES IN
CC RESPONSE TO FUEL LIMITATION AND/OR HYPOXIA. THIS IS A REGULATORY
CC SUBUNIT.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- TISSUE SPECIFICITY: Highly expressed in heart and brain, also
CC found in kidney, white adipose tissue, lung and spleen.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL; X95578; CAA64831.1; -
DR EMBL; U42413; AAC52580.1; -

```

DR PIR; T10759;
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 47 95 CBS 1.
FT DOMAIN 122 176 CBS 2.
FT DOMAIN 136 249 CBS 3.
FT DOMAIN 270 322 CBS 4.
FT CONFLICT 114 114 E -> Q (IN REF. 3).
FT CONFLICT 201 201 A -> P (IN REF. 3).
SQ SEQUENCE 330 AA; 37386 MW; 36031E526C1F1B97 CRC64;

Query Match 15.2%; Score 355.5; DB 1; Length 330;
Best Local Similarity 30.3%; Pred. No. 1.8e-19;
Matches 100; Conservative 67; Mismatches 124; Indels 39; Gaps 11;

QY 121 YLNLHTCYLLPDSGKVIADINLPVKQSFHILHEQIPVAPLWDSFRGQFVGLLSPDLF 180
Db 31 FMKSHRCYDLIPTSSKLVVFDTSLOVKKAFALVNGVRAAPLWDSKQSFVGLITIDF 90
QY 181 ILILRETHGSLNTE-EOLETHITISAWKEAKQTNGRNDOSQWRPOOHLVHATPYESLRD 239
Db 91 INILHRY--YKSALVQIYELEHKEIKETWREVLYQ-----DSFKP---LVCISPNASLED 139
QY 240 IAVKLLONGISTVPVIYSSSDSGSPQLLHLASLSGLKXICICRYKXNCTGNLPILNQPV 299
Db 140 AVSSLRNKIHLRPVIDPSGN-----TLYLTHKRLKFL-KLFTTEFPKPFMSKSL 193
QY 300 SIPLGSGWPEKIGDLNSRPLMLRPNASLSALNMLVQAGVSSIPIVDDNDSLDTYSRSD 359
Db 194 ELIQGIYA-----NIAWVTTTPVVALGIFVQHRVSALPVVDEKGRVVDIYSKED 244
QY 360 ITALAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCQMLRSDPLLKWMERLAN 419
Db 245 VINLAAEKTNNL--DVSVTKALQHRSH-----YFEG--VLKCYLHETLETIINRLVE 293
QY 420 PGVRRVFIIVEAGSKRVEGIISLSDIFKELL 449
Db 294 AEVHRLVVVDEHDV-VKGIIVSLSDILQDLV 322

RESULT 5
AAKG MOUSE
ID AAKG MOUSE STANDARD; PRT; 330 AA.
AC 054950;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPKg).
GN PRKAG1 OR PRKAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21203559; PubMed=11306812;
RA Shamsadin R., Jantsan K., Adham I., Engel W.;
RT "cloning, organisation, chromosomal localization and expression
RT analysis of the mouse Prkag1 gene".
RL Cytogenet. Cell Genet. 92:134-138 (2001).
CC -1- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.
CC -1- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -1- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.

```

```

CC CC -1- SIMILARITY: Contains 4 CBS domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF036535; AAB95475.1; -.
DR MGD; MGI:108411; Prkag1.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 47 95 CBS 1.
FT DOMAIN 122 176 CBS 2.
FT DOMAIN 136 249 CBS 3.
FT DOMAIN 270 322 CBS 4.
SQ SEQUENCE 330 AA; 37554 MW; D504707B83512DDC CRC64;

Query Match 14.9%; Score 350; DB 1; Length 330;
Best Local Similarity 27.8%; Pred. No. 4.7e-19;
Matches 107; Conservative 79; Mismatches 135; Indels 64; Gaps 14;

QY 66 INTLSPSTPGSRMNDVDNENFQRTVTLSDGTVSEGLRVSEAAIOISRCRVSEYINLH 125
Db 1 MESVAEASSPA-----LENEHFQETPE-SNNVS-----YTSPMKSH 35
QY 126 TCYDILPDSCKVIALDINLPVKQSEHILHEQIPVAPLWDSFRGQFVGLLSPDLFILLR 185
Db 36 RYDLIPTSSKLVVFDTSLOVKKAFALVNGVRAAPLWDSKQSFVGLITIDFILLH 95
QY 186 ELETHGSLNTE-EOLETHITISAWKEAKQTNGRNDOSQWRPOOHLVHATPYESLRDIAVKL 244
Db 96 RY--YKSALVQIYELEHKEIKETWREVLYQ-----DSFKP---LVCISPNASLEDVSSL 144
QY 245 LONGISTVPVIYSSSDSGSPQLLHLASLSGLKXICICRYKXNCTGNLPILNQPVCSILP 304
Db 145 IRNKIHLRPVIDPSGN-----TLYLTHKRLKFL-KLFTTEFPKPFMSKSLQELQIG 198
QY 305 SWVPKIGDLNSRPLMLRPNASLSALNMLVQAGVSSIPIVDDNDSLDTYSRSDITALA 364
Db 199 IYA-----NIAWVTTTPVVALGIFVQHRVSALPVVDEKGRVVDIYSKEDVINLA 249
QY 365 KDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCQMLRSDPLLKWMERLANPGVRR 424
Db 250 AEKTYNNL--DVSVTKALQHRSH-----YFEG--VLKCYLHETLETIINRLVEAEVHR 298
QY 425 VFIVEAGSKRVEGIISLSDIFKELL 449
Db 299 LVVVDE-HXXVKGIVSLSDILQDLV 322

RESULT 6
AAKG HUMAN
ID AAKG HUMAN STANDARD; PRT; 464 AA.
AC Q9UGI9; Q9NRL1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)
DE (AMPK gamma3).
GN PRKAG3 OR AMPKG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20164049; PubMed=10698692;

```

QY	296	QPVCSITPLGSSWPKIGDILNSPLMLRPNKASLSSALNMLVQAGVSSIPVDDNDSLLDTY	355
DB	321	RTIQDLIGLGTGFT-----RDLAVLEETAPILTDALDFVDRVSAIPVNVGCGVQVGLY	371
QY	356	SRSDDITALAKQVYTHVRUDENITHQALQGDQANTPFQFFNGQRCQMCILRS-----D	408
DB	372	SFPDVIHLAAQQTYNHLL-----DMSVGEALR-----QR-TLCLEGLVLSQCPHE	413
QY	409	PILKWERLHAGVRRVFIVEAGSKRVEGIILSLSDIFKELL	449
DB	414	SLGEVIDRIAREQVHRVLVWDE-TQHLIGVWSLSDILQALV	453
RESULT 7			
AAKL_PIG	AAKL_PIG	STANDARD; PRT; 514 AA.	
ID	Q9MYP4;		
DC	16-OCT-2001 (Rel. 40, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)		
DE	(AMPK gamma3).		
GN	PRKAG3.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT RN(-) GLN-250.		
RC	TISSUE=Skeletal muscle;		
RC	MEDLINE=20280150; PubMed=10818001;		
RA	Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,		
RA	Rogel-Gaillard C., Paul S., Iannucci N., Rask L., Ronne H.,		
RA	Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,		
RA	Andersson L.;		
RT	"A mutation in PRKAG3 associated with excess glycogen content in pig		
RT	skeletal muscle.";		
RL	Science 288:1248-1251(2000).		
RL	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	TISSUE=Skeletal muscle;		
RA	Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,		
RA	Rogel-Gaillard C., Paul S., Gellin J., Lundstrom K., Reinsch N.,		
RA	Kalm E., Le Roy P., Chardon P., Andersson L.;		
RA	Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.		
CC	-!- FUNCTION: AMPK is responsible for the regulation of fatty acid		
CC	synthesis by phosphorylation of acetyl-CoA carboxylase. Also		
CC	regulates cholesterol synthesis via phosphorylation and		
CC	inactivation of hydroxymethylglutaryl-CoA reductase and hormone-		
CC	sensitive lipase. This is a regulatory subunit. It may play a role		
CC	in the regulation of energy metabolism in skeletal muscle.		
CC	-!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a		
CC	gamma non-catalytic regulatory subunits.		
CC	-!- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=2;		
CC	IsoId=Q9MYP4-1; Sequence=Displayed;		
CC	Name=1;		
CC	IsoId=Q9MYP4-2; Sequence=VSP_008059;		
CC	-!- TISSUE SPECIFICITY: MUSCLE.		
CC	-!- DISEASE: Defects in PRKAG3 re the cause of the RN- phenotype which		
CC	is associated with excess glycogen content (about 70%) in skeletal		
CC	muscle. This mutation originated in the Hampshire breed pigs and		
CC	has beneficial effects on meat content but detrimental effects on		
CC	processing yield. Thus, this mutation is of considerable economic		
CC	significance in the pig breeding industry.		
CC	-!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma		
CC	subunit family.		
CC	-!- SIMILARITY: Contains 4 CBS domains.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	at the European Bioinformatics Institute. There are no restrictions on its		


```
QY 235 ESURDRAVXKLQNGISTVPVIVYSSSDGSPFQLHLHSLGILKCI--CR---YFKNST 288
DB 130 RPLFEACLMLESRSGRIPLI-DQDEETHREIVSVLTQYRIILKFPVALNCRHFLKPI 188
QY 289 GNLPIILQPCVCSIPLGSWPKIGDLSRPLMLRPNASLSALNMLVQAGVSSIPVDDN 348
DB 189 GELNIIQD-----NMKSCQMTTP---VIDVQLTQGRVSSVPIIDEN 229
QY 349 LSLIDTYSRSDITAKADKYVTEVRLDEMTHQALQIQDANTPFGFNGRCOMCLUSD 408
DB 230 GYLINVEADVILGKGIYNDLSL---SVGELMRSSDD-----FEG--VYTCKND 278
QY 409 ELKVMERLANPGVRRVFIYEASGRVEGIIISDIEFKLL 449
DB 279 KLSTINDIRKARVHRFFVVD-DVGRIVGVLTLSDLIKYIL 318

RESULT 9
YL28-SCHPO STANDARD; PRT; 334 AA.
ID YL28 SCHPO
AC Q10343; Q9UTJ1;
DT 01-OCT-1996 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein GL556.08c in chromosome 1.
GN SPAC1556.08C OR SPAC1F12.01C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton C., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward T., Volkart G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Welter J., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Reinhardt R., Pohl T.M.,
RA Borzym K., Langer I., Beck A., Jehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purcell B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; Z69944; CAA93805.1; --
DR EMBL; ALJ32984; CAB61219.1; --
DR GeneDB SPombe; SPAC1556.08c; --
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 33 81 CBS 1.
FT DOMAIN 118 170 CBS 2.
FT DOMAIN 194 241 CBS 3.
FT DOMAIN 266 314 CBS 4.
SQ SEQUENCE 334 AA; 37427 MW; 8D32CC6CE53F7916 CRC64;

Query Match 12.38; Score 289; DB 1; Length 334;
Best Local Similarity 26.0%; Pred. No. 1.9e-14; Indels 36; Gaps 12;
Matches 87; Conservative 74; Mismatches 137;

QY 118 VSEYLNLHTICVDLLPDSGKVIADINLFPVQSFILHEQGIPEVAPLWDSFPGQFVGLLSP 177
DB 15 IQAFIRSETSYDVLPTSFRLIVFDVTLFVKTSLSLLTLNNIVSAPLWDSSEANKFAGLLTM 74
QY 178 LDFILILRELETHGSLNTE--EQLETHYISAMKEAKRQTNGRNDSQWRPQOHLVHATPYE 235
DB 75 ADFVNVIK-YYIQSSFPFEALAEIDKFRILGLREVERKIGA-----IPPTIYYVH--PMH 126
QY 236 SLRDIYVKLQNGISTVPVIVYSSSDGSPFQLHLHSLGILKCIYFKNSTGMLPLIN 295
DB 127 SLMDACLAMSKSRARRIPLIDVDGETS-EMIVSVLTQYRIILKFIWMCKETA---MLR 181
QY 296 QPVCSIPLGSWPKIGDLSRPLMLRPNASLSALNMLVQAGVSSIPVDDNDSLLDTY 355
DB 182 VPLNQMTIGTW-----SNLATASMETKVYDVIVKMLAEKNISAVFIVNSEGLINNVY 232
QY 356 SRSDITALAKDKVYTHVRLDEMTHQALQIQDANTPFGFNGRCOMCLASDILLKYWE 415
DB 233 ESDVWMLHIDQGYSNL---DLSGEAL-LKRPAN-----FDG--VHTCRATDRLDGIFD 281
QY 416 RLAMPGVRRVFIYEASGRVEGIIISDIEFKLL 449
DB 282 AIKHSRVHRLFPVVDENLK-LEGILSLADILNYII 314

RESULT 10
SNF4_KJULA STANDARD; PRT; 328 AA.
ID SNF4_KJULA
AC Q9B869;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nuclear protein SNF4.
GN SNF4.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RA Tomasini L., Ferrero I., Goffrini P.;
RT "Molecular characterization of KLSNF4 gene."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This protein causes expression of glucose-repressible
CC genes upon glucose deprivation. It interacts and has functional
CC relationship to the protein-kinase SNF1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity)
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```


CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U42411; RAC52579.1; --
DR EMBL; X95577; CAA64830.1; --
DR InterPro; IPR006828; AMPKBI.
DR Pfam; PF04739; AMPKBI; 1.
KW Fatty acid biosynthesis; Phosphorylation; Myristate; Lipoprotein.
FT INIT_MET 0 0
FT LIPID 1 1 N-myristoyl glycine.
FT MOD_RES 23 23 PHOSPHORYLATION (AUTO-).
FT MOD_RES 24 24 PHOSPHORYLATION (AUTO-).
FT MOD_RES 107 107 PHOSPHORYLATION (AUTO-).
FT MOD_RES 181 181 PHOSPHORYLATION.
FT MOD_RES 25 25 G -> E (IN REF. 1).
FT CONFLICT 51 51 M -> I (IN REF. 3; AA SEQUENCE).
FT CONFLICT 51 51 M -> I (IN REF. 3; AA SEQUENCE).
SQ SEQUENCE 269 AA; 30263 MW; 6245087E57E581E1 CRC64;

Query Match 5.3%; Score 124; DB 1; Length 269;
Best Local Similarity 32.5%; Pred. No. 0.036;
Matches 27; Conservative 14; Mismatches 26; Indels 16; Gaps 2;

QY 9 PIEGCPVFOAICSLSPGIIHEVYKPFVGDGEWRHDSRQPTISGEFIVNTLYL----- 59
DB 103 PLTRSQNNFVALDLPGEHQYKPFVGDGQWTHDPSEPIVTSQLGIVNNIIQVKKTDFFVF 162

QY 60 -----TREYNQINTLSPSTPG 76
DB 163 DALMVDQKSCDSVLSLSS-SPPG 184

RESULT 13
AAKB HUMAN
ID AAKB HUMAN STANDARD; PRT; 269 AA.
AC Q9Y478; Q9Y478; Q9Y478; Q9Y478; Q9Y478; Q9Y478; Q9Y478; Q9Y478;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, beta-1 subunit (AMPK beta-1 chain)
DE (AMPKb).
GN PRKAB1 OR AMPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Carling D.;
RT "Non-catalytic beta and gamma subunits isoforms of the AMP-activated
protein kinase.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=97367941; PubMed=9224708;
RA Stapleton D., Woollett E., Mitchell K., Nicholl J.K.,
RA Fernandez C.S., Mitchell B.J., Witters L.A., Power D.A.,
RA Sutherland G.R., Kemp B.E.;
RA "AMP-activated protein kinase isoenzyme family: subunit structure and
chromosomal location.";
RT FEBS Lett. 409:452-456(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Yamagata K., Oda N., Furuta H., Vaxillaire M., Southam L., Boriraj V.,
RA Chen X., Oda Y., Takeda J., Yamada S., Nishigori H., Lebeau M.M.,
RA Lathrop M., Cox R.D., Bell G.I.;
RA "Transcription map of the 5CM region surrounding the hepatocyte
nuclear factor-1a/MODY3 gene on chromosome 12.";
RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]

RP SEQUENCE FROM N.A.
 RA Wang X., Xu L., Tu Q.;
 RT "Cloning and expression of the complete mRNA coding human AMP-
 RL activated protein kinase."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Connell M., Goela D., Harper M.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Lung, and Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Jedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska J., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16839-16903(2002).
 CC -!- FUNCTION: AMPK is responsible for the regulation of fatty acid
 CC synthesis by phosphorylation of acetyl-CoA carboxylase. Also
 CC regulates cholesterol synthesis via phosphorylation and
 CC inactivation of hydroxymethylglutaryl-CoA reductase and hormone-
 CC sensitive lipase. This is a regulatory subunit, may be a positive
 CC regulator of AMPK activity. It may also serve as an adaptor
 CC molecule for the catalytic alpha-subunit.
 CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
 CC gamma non-catalytic regulatory subunits.
 CC -!- PTM: Phosphorylated (By similarity).
 CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta
 CC subunit family.
 CC -!- CAUTION: Ref.4 sequence differs from that shown due to a
 CC frameshift in position 244.
 CC -!- CAUTION: Ref.5 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
 CC or send an email to license@isb-sib.ch).
 CC
 DR ENBL; AJ224515; CAA12024.1; -;
 DR ENBL; Y12556; CAA73146.1; -;
 DR ENBL; U83994; AAD09237.1; -;
 DR ENBL; U87276; RAD00625.1; -;
 DR ENBL; U87271; AAD00625.1; JOINED.
 DR ENBL; U87272; RAD00625.1; JOINED.
 DR ENBL; U87273; RAD00625.1; JOINED.
 DR ENBL; U87274; RAD00625.1; JOINED.
 DR ENBL; U87275; RAD00625.1; JOINED.
 DR ENBL; AF022116; AAC98897.1; ALT FRAME.
 DR ENBL; AC002563; AAB71326.1; ALT_SEQ.
 DR ENBL; BC001007; AAH01007.1; -;
 DR ENBL; BC001056; AAH01056.1; -;
 DR ENBL; BC001823; AAH01823.1; -;
 DR ENBL; BC017671; AAH17671.1; -;

DR PIR; T09514; T09514.
 DR Genew; HGNC:9378; PRKAB1.
 DR MIM; 602740; -;
 DR GO; GO:0007185; P:signal transduction; TAS.
 DR InterPro; IPR006828; AMPKBI.
 DR Pfam; PF04739; AMPKBI; 1.
 DR KW Fatty acid biosynthesis; Phosphorylation; Myristate; Lipoprotein.
 DR INIT MET 0 0
 DR BY SIMILARITY.
 FT LIPID 1 1 N-myristoyl glycine (By similarity).
 FT MOD RES 23 23 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 24 24 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 107 107 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 181 181 PHOSPHORYLATION (BY SIMILARITY).
 FT CONFLICT 9 9 A -> G (IN REF. 2 AND 4).
 FT CONFLICT 14 14 G -> A (IN REF. 1).
 FT CONFLICT 19 19 P -> A (IN REF. 2 AND 4).
 FT CONFLICT 21 21 R -> K (IN REF. 3).
 FT CONFLICT 55 55 E -> Y (IN REF. 3).
 SQ SEQUENCE 269 AA; F08BCF39D5AAF861 CRC64;
 Query Match 5.3%; Score 123; DB 1; Length 269;
 Best Local Similarity 34.7%; Pred. No. 0.043; Indels 16; Gaps 2;
 Matches 26; Conservative 13; Mismatches 20;
 QY 17 FQALCSLPGHEHYKFFVDGSEWRHDERQPTISGFIVNTLYL-----TR 61
 DB 111 FVALDLPEGEHQYKFFVDGQWTHDPSEPIVTSQLTGVTNNIIQVKKTKDFEVDALMVDQ 170
 QY 62 EYNQINTLSSTPTG 76
 DB 171 KCDVSELS-SPFG 184
 RESULT 14
 AKAC RAT
 ID AKAC RAT STANDARD; PRT; 271 AA.
 AC Q9QZH4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 5'-AMP-activated protein kinase, beta-2 subunit (AMPK beta-2 chain).
 GN PRKAB2.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Skeletal muscle;
 RX MEDLINE=20012956; PubMed=10544261;
 RA Chen Z., Heierhorst J., Mann R.J., Mitchell K.I., Mitchell B.J.,
 RA Witters L.A., Lynch G.S., Kemp B.E., Stapleton D.;
 RT "Expression of the AMP-activated protein kinase beta1 and beta2
 RT subunits in skeletal muscle."
 RL FEBS Lett. 460:343-348(1999).
 CC -!- FUNCTION: AMPK is responsible for the regulation of fatty acid
 CC synthesis by phosphorylation of acetyl-CoA carboxylase. Also
 CC regulates cholesterol synthesis via phosphorylation and
 CC inactivation of hydroxymethylglutaryl-CoA reductase and hormone-
 CC sensitive lipase. This is a regulatory subunit, may be a positive
 CC regulator of AMPK activity. It may also serve as an adaptor
 CC molecule for the catalytic alpha-subunit.
 CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
 CC gamma non-catalytic regulatory subunits.
 CC -!- PTM: Phosphorylated when associated with the catalytic subunit.
 CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta
 CC subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF182717; AA01293.1; -
CC InterPro; IPR006828; AMPKXI.
CC Pfam; PF04733; AMPKXI; 1.
KW Fatty acid biosynthesis; Phosphorylation.
FT MOD_RES 183 183 PHOSPHORYLATION (PARTIAL).
SQ SEQUENCE 271 AA; 30227 MW; D881451A03287BEA CRC64;

Query Match 5.2%; Score 122.5; DB 1; Length 271;
Best Local Similarity 31.5%; Pred. No. 0.047;
Matches 29; Conservative 13; Mismatches 31; Indels 19; Gaps 3;

Qy 17 FOAICSLSPGHEHYKFFVDGWRHDEROPTSGRFGIVNTLYLTK--SYNQINTL----- 69
Db 111 FVALIDLPEGEHQYKFFVDGQWHDPSFVVTSQLGTINNLHVKKSDFEVFDALKLDSM 170

Qy 70 -----SSSTP-GSRMMDVDNENFQ 89
Db 171 ESSTCRDLSSPPGPGYQEMVFRSERFK 202

RESULT 15
RAKC_HUMAN
ID -RAKC_HUMAN STANDARD; PRT; 272 AA.
AC O43741;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5'-AMP-activated protein kinase, beta-2 subunit (AMPK beta-2 chain).
GN PRKAB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominalae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98241616; PubMed=9575201;
RA Thornton C., Scowden M.A., Carling D.;
RT "Identification of a novel AMP-activated protein kinase beta subunit
RT isoform that is highly expressed in skeletal muscle.";
RL J. Biol. Chem. 273:12443-12450(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22379004; PubMed=12490143;
RA Prochazka M., Farock V.S., Ossowski V., Wolford J.K., Bogardus C.;
RT "Variant screening of PRKAB2, a type 2 diabetes mellitus
RT susceptibility candidate gene on 1q in Pima Indians.";
RL Mol. Cell. Probes 16:421-427(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Pancreas;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska J., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
```

```
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: AMPK is responsible for the regulation of fatty acid
synthesis by phosphorylation of acetyl-CoA carboxylase. Also
regulates cholesterol synthesis via phosphorylation and
inactivation of hydroxymethylglutaryl-CoA reductase and hormone-
sensitive lipase. This is a regulatory subunit, may be a positive
regulator of AMPK activity. It may also serve as an adaptor
molecule for the catalytic alpha-subunit.
-!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
gamma non-catalytic regulatory subunits.
-!- PTM: Phosphorylated when associated with the catalytic subunit.
-!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta
subunit family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AJ224538; CAAL2030.1; -
DR EMBL; AF504543; AAM74153.1; -
DR EMBL; AF504538; AAM74153.1; JOINED.
DR EMBL; AF504539; AAM74153.1; JOINED.
DR EMBL; AF504540; AAM74153.1; JOINED.
DR EMBL; AF504541; AAM74153.1; JOINED.
DR EMBL; AF504542; AAM74153.1; JOINED.
DR EMBL; BC053610; AAH53610.1; -
DR EMBL; HGNC:9379; PRKAB2.
DR MIM; 602741; -
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR006828; AMPKXI.
DR Pfam; PF04739; AMPKXI; 1.
KW Fatty acid biosynthesis; Phosphorylation.
FT MOD_RES 183 183 PHOSPHORYLATION (PARTIAL) (BY
SIMILARITY).
SQ SEQUENCE 272 AA; 30302 MW; 42B23BD70B92519C CRC64;

Query Match 5.2%; Score 121; DB 1; Length 272;
Best Local Similarity 48.8%; Pred. No. 0.061;
Matches 20; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 17 FOAICSLSPGHEHYKFFVDGWRHDEROPTSGRFGIVNTL 57
Db 112 FVALIDLPEGEHQYKFFVDGQWHDPSFVVTSQLGTINNL 152

Search completed: July 7, 2004, 17:51:20
Job time : 19 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:49:44 ; Search time 52 Seconds

(without alignments)

2699.788 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREHLPMSPIECPTVFQAI.....SXRVEGILSLDIEKFLLSL 451

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/USC6_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/USC6_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	95.3	514	12	US-10-425-114-38422 Sequence 38422, A
2	2230	95.2	514	12	US-10-425-114-64906 Sequence 64906, A
3	2178	93.0	423	12	US-10-425-114-38456 Sequence 38456, A
4	1944	83.0	723	16	US-10-437-963-105356 Sequence 105356, A
5	1802	76.9	365	12	US-10-425-114-69679 Sequence 69679, A
6	1527.5	65.2	493	12	US-10-424-599-187700 Sequence 187700, A
7	1492	63.7	302	12	US-10-425-114-58697 Sequence 58697, A
8	1487	63.5	461	12	US-10-424-599-225531 Sequence 225531, A
9	1325	56.6	368	12	US-10-425-114-54796 Sequence 54796, A
10	923.5	39.4	263	12	US-10-424-599-142943 Sequence 142943, A
11	920	39.3	299	12	US-10-424-599-142947 Sequence 142947, A
12	908.5	38.8	2248	16	US-10-437-963-193213 Sequence 193213, A
13	820.5	35.0	451	16	US-10-437-963-168583 Sequence 168583, A
14	778	33.2	166	16	US-10-437-963-125316 Sequence 125316, A
15	761	32.5	477	12	US-10-425-114-50202 Sequence 50202, A

16	741.5	31.7	230	12	US-10-424-599-142942	Sequence 142942, A
17	580.5	24.8	232	12	US-10-425-114-38434	Sequence 38434, A
18	425.5	18.2	255	12	US-10-425-114-57823	Sequence 57823, A
19	373	15.9	328	10	US-09-769-970-7	Sequence 7, Appli
20	367	15.7	331	10	US-09-769-970-21	Sequence 21, Appl
21	359	15.3	344	9	US-09-925-297-461	Sequence 461, Appl
22	359	15.3	1207	13	US-10-108-605-71	Sequence 71, Appl
23	356	15.2	193	12	US-10-424-599-187709	Sequence 187709, A
24	340	14.5	489	9	US-09-826-581-6	Sequence 6, Appli
25	340	14.5	489	16	US-10-473-670-15	Sequence 15, Appl
26	340	14.5	489	16	US-10-705-137-6	Sequence 6, Appli
27	335.5	14.3	464	12	US-09-950-022-8	Sequence 8, Appli
28	335.5	14.3	464	12	US-09-950-022-10	Sequence 10, Appl
29	334.5	14.3	464	12	US-09-950-022-2	Sequence 2, Appli
30	334.5	14.3	464	12	US-09-950-022-4	Sequence 4, Appli
31	334.5	14.3	464	12	US-09-950-022-6	Sequence 6, Appli
32	297.5	12.7	120	12	US-10-424-599-148301	Sequence 148301, A
33	231	9.9	136	12	US-10-424-599-187699	Sequence 187699, A
34	173	7.4	149	15	US-10-264-237-1653	Sequence 1653, Ap
35	153	6.5	92	12	US-10-424-599-197913	Sequence 197913, A
36	151	6.4	57	12	US-10-424-599-210108	Sequence 210108, A
37	133	5.7	83	12	US-10-424-599-187707	Sequence 187707, A
38	130.5	5.6	180	14	US-10-106-698-6572	Sequence 6572, Ap
39	115.5	4.9	443	12	US-10-424-599-223551	Sequence 223551, A
40	111.5	4.8	181	12	US-10-276-774-1611	Sequence 1611, Ap
41	110.5	4.7	446	12	US-10-282-122A-69591	Sequence 69591, A
42	108.5	4.6	425	16	US-10-437-963-161228	Sequence 161228, A
43	107.5	4.6	4910	12	US-10-346-863-21	Sequence 21, Appl
44	107	4.6	476	12	US-10-282-122A-60388	Sequence 60388, A
45	106.5	4.5	1274	16	US-10-437-963-196165	Sequence 196165, A

ALIGNMENTS

RESULT 1

US-10-425-114-38422
; Sequence 38422, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38422
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700172325_FLI.pep
US-10-425-114-38422

Query Match	95.3%;	Score	2231;	DB	12;	Length	514;
Best Local Similarity	95.5%;	Pred. No.	4.8e-213;				
Matches	429;	Conservative	8;	Mismatches	12;	Indels	0;
Qy	3	EHLPMSPICGCTVFQAI	CSLSPGIEHYKFFVDCGWRHDERQPTISGEFGVNTLYLTRE	62			
		:	:	:	:	:	:
Db	66	EHLPMSPICGCTVFQAI	CSLSPGIEHYKFFVDCGWRHDERQPTISGEFGVNTLYLTRE	125			
		:	:	:	:	:	:
Qy	63	YNQINTLSPSTPGSRMMDVDNENFQRTVTLSDGTWSEGLRVSEAAIQISRCRVSEYL	122				
		:	:	:	:	:	:
Db	126	FNQINALLNPTPGSRMMDVDNENFQRTVTLSDGTIPGIVRVSEAAIQISRCRVSEYL	185				
		:	:	:	:	:	:
Qy	123	NLHTCYDLLPDSKGVIALDINLPVQSFHILHEGQIPVAPLWDSFRQGFVGLLSPLDIL	182				
		:	:	:	:	:	:

```

Db      186 NLHTCYDLPDSKVALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPLDFIL 245
Qy      183 ILRELETHGSLNTEQLETHTISAWKEAKQTGRNDQSWRPQOHLVHATPYESLRDIIV 242
Db      246 ILRELETHGSLNTEQLETHTISAWKEAKQTGRNDQSWRPQOHLVHATPYESLRDIIV 305
Qy      243 KLLONGISTVPVIYSSSDGSPFQLHLASLSGILKICRYFNKSTGNLPILOPVCISIP 302
Db      306 KLLNDISTVPVIYSSSDGSPFQLHLASLSGILKICRYFNKSTGNLPILOPVCISIP 365
Qy      303 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSPIVDDNDSLLDTYSRSDITA 362
Db      366 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSPIVDDNDSLLDTYSRSDITA 425
Qy      363 LAKDKVYTHVRLDEMTIHOALQOLQDANTPFGFNQRCQMLRSDPDLKXMERLANPGV 422
Db      426 LAKDKVYTHVRLDEMTIHOALQOLQDANTPFGFNQRCQMLRSDPDLKXMERLANPGV 485
Qy      423 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 451
Db      486 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 514

```

RESULT 2

```

US-10-425-114-64906
; Sequence 64906, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73:28
; SEQ ID NO 64906
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4729-C18-A12_FLI.pep
US-10-425-114-64906

```

```

Query Match      95.2%; Score 2230; DB 12; Length 514;
Best Local Similarity 95.3%; Pred. No. 6e-213;
Matches 428; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

Qy      3 EHLPMSPIEGCTVFQAICSLSPGIEHYKFPVDEGWRHDERQPTISGEFIVNTLYLTRE 62
Db      66 EHLPMSPVEGCTVFQAICSLSPGIEHYKFPVDEGWRHDERQPTISGEFIVNTLYLTRE 125
Qy      63 YNQINTLSSPSTPGSRNMVDNENFQRTVTLSDGTVSCTLRVSEAAIQISRCRVSEYL 122
Db      126 FNQINALLNFTPGSRNMVDNENFQRTVTLSDGTIPBGTVRVSEAAIQISRCRVSEYL 185
Qy      123 NLHTCYDLPDSGKVIALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPLDFIL 182
Db      186 NLHTCYDLPDSGKVIALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPLDFIL 245
Qy      183 ILRELETHGSLNTEQLETHTISAWKEAKQTGRNDQSWRPQOHLVHATPYESLRDIIV 242
Db      246 ILRELETHGSLNTEQLETHTISAWKEAKQTGRNDQSWRPQOHLVHATPYESLRDIIV 305
Qy      243 KLLONGISTVPVIYSSSDGSPFQLHLASLSGILKICRYFNKSTGNLPILOPVCISIP 302
Db      306 KLLNDISTVPVIYSSSDGSPFQLHLASLSGILKICRYFNKSTGNLPILOPVCISIP 365

```

```

Qy      303 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSPIVDDNDSLLDTYSRSDITA 362
Db      366 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSPIVDDNDSLLDTYSRSDITA 425
Qy      363 LAKDKVYTHVRLDEMTIHOALQOLQDANTPFGFNQRCQMLRSDPDLKXMERLANPGV 422
Db      426 LAKDKVYTHVRLDEMTIHOALQOLQDANTPFGFNQRCQMLRSDPDLKXMERLANPGV 485
Qy      423 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 451
Db      486 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 514

RESULT 3
US-10-425-114-38456
; Sequence 38456, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73:28
; SEQ ID NO 38456
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700342215_FLI.pep
US-10-425-114-38456

```

```

Query Match      93.0%; Score 2178; DB 12; Length 423;
Best Local Similarity 99.5%; Pred. No. 6.7e-208;
Matches 421; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      29 EYKFFVDGWRHDERQPTISGEFIVNTLYLTRYNQINTLSSPSTPGSRNMVDNENF 88
Db      61 EYKFFVDGWRHDERQPTISGEFIVNTLYLTRYNQINTLSSPSTPGSRNMVDNENF 60
Qy      89 QRTVTLSDGTVSCTLRVSEAAIQISRCRVSEYKMLHTCYDLPDSGKVIALDINLPVKQ 148
Db      61 QRTVTLSDGTVSCTLRVSEAAIQISRCRVSEYKMLHTCYDLPDSGKVIALDINLPVKQ 120
Qy      149 SFHILHEQGPVAPLWDSFRGQFVGLLSPLDFILILRELETHGSLNTEQLETHTISAWK 208
Db      121 SFHILHEQGPVAPLWDSFRGQFVGLLSPLDFILILRELETHGSLNTEQLETHTISAWK 180
Qy      209 EAKQRTGRNDQSWRPQOHLVHATPYESLRDIIVKLLONGISTVPVIYSSSDGSPFOLL 268
Db      181 EAKQRTGRNDQSWRPQOHLVHATPYESLRDIIVKLLONGISTVPVIYSSSDGSPFOLL 240
Qy      269 HLASLSGILKICRYFNKSTGNLPILOPVCISIPGSMVPKIGDLSRPLAMLRPNASLS 328
Db      241 HLASLSGILKICRYFNKSTGNLPILOPVCISIPGSMVPKIGDLSRPLAMLRPNASLS 300
Qy      329 SALNMLVQAGVSSPIVDDNDSLLDTYSRSDITALAKDKVYTHVRLDEMTIHOALQOLQD 388
Db      301 SALNMLVQAGVSSPIVDDNDSLLDTYSRSDITALAKDKVYTHVRLDEMTIHOALQOLQD 360
Qy      389 ANTFFGFFNGRCQMLRSDPDLKXMERLANPGVRRVFRVIVEAGSKRVEGIISLSDIFKFL 448
Db      361 ANTFFGFFNGRCQMLRSDPDLKXMERLANPGVRRVFRVIVEAGSKRVEGIISLSDIFKFL 420
Qy      449 LSL 451
Db      421 LSL 423

```


RESULT 4

```

RES001  *
US-10-437-963-105356
; Sequence 105356, Application US/10437963
; Publication No. US20040123343A1
;
; GENERAL INFORMATION:
;
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbaruk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Mols
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 105356
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Oryza sativa
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: PAT MFT4530
US-10-437-963-105356

```

5 2.11544

RESULT 5
US-10-425-114-69679
; Sequence 69679, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69679
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMOL17174G09_FLI.pep
US-10-425-114-69679

```

DEC 11 1966

```

RESULT 6
US-10-424-599-187700
: Sequence 187700, Application US/10424599
: Publication No. US20040031072A1
:
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 187700
: LENGTH: 493
: TYPE: PRT
: ORGANISM: Glycine max
:
: FEATURE:
:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_140506C.1.psp
: US-10-424-599-187700

```

QY 3 EHLPMSPTEGCTVFQALCSLSPGHEHYKFFVDEWHRDERQCTISGEFGIVNTL-YLTR 61
 Db 46 ELLQMSPEVGCTVFQVHSHLLPGYHQYKFFVDEWHRDERQCTISGEFGIVNTL-YLTR 105
 QY 62 EYNOINTLSSPSTPGSRMNDVDNENFQRTVTLSGTVSEGTLRVSEAAIQISRCRVSEY 121
 Db 106 DPNITPVL--PPDVASGNSMDVDNDARFARXARLTDGILSEVLPRISTDVDQISQRISAF 163
 QY 122 LNLHTCYDLLPDSQKVALDINLVKQSPHILHGGQIPVAPLWDSFRGQFVGLLSPDLFI 181
 Db 164 LSSHVAVELLPESQKVALDINLVKQSPHILHGGQIPVAPLWDSFRGQFVGLLSPDLFI 223
 QY 182 LILRELTGHSNLTETQLETHITSAWKEAK-----RQNGRNDNQRRQQHVLVHATPYESL 237
 Db 224 LILRELTGHSNLTETQLETHITSAWKEAK-----RQNGRNDNQRRQQHVLVHATPYESL 280
 QY 238 RDIKVLKQNGISVTPVIVSYSSSDGSPQLHLHSLGILKICRYFKNSTGNLPILNQ 297
 Db 281 KDIAMKILQXSVTPVIVSYSSSDGSPQLHLHSLGILKICRYFKNSTGNLPILNQ 340
 QY 298 VCSIPGSGWPKIGMNSRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 357
 Db 341 ICAIPVGTWPKIGESNRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 400
 QY 358 SDITLAKDKVYTHVRLEDEMTIHQALQIGODANTPPGFNGQRCOMCLRSPLIKVMERL 417
 Db 401 SDITLAKDKVYTHVRLEDEMTIHQALQIGODANTPPGFNGQRCOMCLRSPLIKVMERL 459
 QY 418 ANPGVRVFIIVEAGSKRVEGIIISLSDIFKFL 449
 Db 460 ANPGVRVFIIVEAGSKRVEGIIISLSDIFKFL 491

RESULT 7

US-10-425-114-58697
 ; Sequence 58697, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53333)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 58697
 ; LENGTH: 302
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700575557_FJ1.pep
 US-10-425-114-58697

Query Match 63.7%; Score 1492; DB 12; Length 302;
 Best Local Similarity 96.7%; Pred. No. 9.9e-140;
 Matches 292; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 150 PHILHEQIPVAPLWDSFRGQFVGLLSPDLILRELETHGNSLTETQLETHITSAWKE 209
 Db 1 PHILHEQIPVAPLWDSFRGQFVGLLSPDLILRELETHGNSLTETQLETHITSAWKE 60
 QY 210 AKRQTCGRNDQWRAHQLVHATPYESLRIKVLQNGISTVPIVYSSSDGSGFPQLLH 269
 Db 61 AKRQTCGRNDQWRAHQLVHATPYESLRIKVLQNGISTVPIVYSSSDGSGFPQLLH 120
 QY 270 LASLSGILKICRYFKNSTGNLPILNQVCSIPGSGWPKIGMNSRPLAMLRPNASLS 329

Db 121 LASLSGILKICRYFKNSTGNLPILNQVCSIPGSGWPKIGMNSRPLAMLRPNASLS 180
 QY 330 ALNMLVQAGVSSIPVDDNDLSDTYSRSDITLAKDKVYTHVRLEDEMTIHQALQIGQDA 389
 Db 181 ALNMLVQAGVSSIPVDDNDLSDTYSRSDITLAKDKVYTHVRLEDEMTIHQALQIGQDA 240
 QY 390 NTPGFGNGQRCOMCLRSPLIKVMERLANPGVRVFIIVEAGSKRVEGIIISLSDIFKFL 449
 Db 241 NTPGFGNGQRCOMCLRSPLIKVMERLANPGVRVFIIVEAGSKRVEGIIISLSDIFKFL 300
 QY 450 SL 451
 Db 301 SL 302

RESULT 8
 US-10-424-599-222531
 ; Sequence 222531, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J
 ; APPLICANT: Kovalic, David K
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 222531
 ; LENGTH: 461
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_42975C.1.pep
 US-10-424-599-222531

Query Match 63.5%; Score 1487; DB 12; Length 461;
 Best Local Similarity 63.6%; Pred. No. 6.2e-139;
 Matches 287; Conservative 65; Mismatches 75; Indels 24; Gaps 4;

QY 4 HLPMSPIEGCTVFQALCSLSPGHEHYKFFVDEWHRDERQCTISGEFGIVNTLTYREY 63
 Db 29 HLPMSPIEGCTVFQALCSLSPGHEHYKFFVDEWHRDERQCTISGEFGIVNTLTYREY 88
 QY 64 NCINTLSSPSTPGSRMNDVDNENFQRTVTLSGTVSEGTLRVSEAAIQISRCRVSEYN 123
 Db 89 DILPSILNTETPG-RSHMEVDNMEANPRMTVSD-----LEVSRRHSISVFLS 133
 QY 124 LHTCYDLLPDSQKVALDINLVKQSPHILHGGQIPVAPLWDSFRGQFVGLLSPDLFI 183
 Db 134 THTSYDLLPESQKVALDINLVKQSPHILHGGQIPVAPLWDSFRGQFVGLLSPDLFI 193
 QY 184 LRELETHGNSLTETQLETHITSAWKEAKRQNGRNDQ-----WRPQOHLVHATPYESL 238
 Db 194 LKELGNHSSNLTQLETHITSAWKEAKRQNGRNDQ-----FVHAGEHECLK 249
 QY 239 DIAVKLLQNGISTVPIVYSSSDGSGFPQLLHSLGILKICRYFKNSTGNLPILNQPV 298
 Db 250 DIAVKLLQNGISTVPIVYSSSDGSGFPQLLHSLGILKICRYFKNSTGNLPILNQPV 309
 QY 299 CSIPGSGWPKIGMNSRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 358
 Db 310 ASIPLGTWMPVPGNGRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 369
 QY 359 DITLAKDKVYTHVRLEDEMTIHQALQIGQDANTPPGFNGQRCOMCLRSPLIKVMERLA 418
 Db 370 DITLAKDKVYTHVRLEDEMTIHQALQIGQDANTPPGFNGQRCOMCLRSPLIKVMERLA 429
 QY 419 NPGVRVFIIVEAGSKRVEGIIISLSDIFKFL 449
 Db 430 NPGVRVFIIVEAGSKRVEGIIISLSDIFKFL 460

RESULT 9

US-10-425-114-54796
; Sequence 54796, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jigdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54796
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-QWFLMINSOY057G02_FLI.pep
US-10-425-114-54796

Query Match 56.6%; Score 1325; DB 12; Length 368;
Best Local Similarity 69.2%; Pred. No. 6.1e-123;
Matches 256; Conservative 52; Mismatches 54; Indels 8; Gaps 3;
QY 84 DNEPOTVTLSDGTSEGLRSEAAIQISRCRWSEYLNLCYDILLPSDGGKVALDIN 143
DB 1 DDAFAFMARLTDTLSEVLPRISDVTQVLSRQRIASFSSHTAYELLPESGKVALD 60
QY 144 LPVKQSFILHEQGIPVAPLWDSFRGQVGLSPDLFILRELETHGSLNLETHET 203
DB 61 LPVKQAFHILHEQGVFMAPLWDFCKGFVGLSASDFILRELGNGHSLNLETHET 120
QY 204 ISAWKEAK----RQTGRNDSQWRPQOHLVHATPYESLRDIAVKLQNGISTVPIYSS 259
DB 121 ISAWKEGKSYLNQNGHGTAFSR--CFIHAGPYDLKDIAMKILQKEVSTVPIHSS 177
QY 260 SDGSPQLHLASLGLKICRYFKNSTGNLPINOPVCSIPGLSVPKIGDINSRPLA 319
DB 178 EDASFPQLHLASLGLKICRYFHCSSSLPVQLPICAIPWGVWPKIGESNRPLA 237
QY 320 MLRPNASLSALNMLVOAGVSSPIVDDNDSLLDTYSRSDITALAKKXVTHVRLD 379
DB 238 MLRPTASLASALNMLVOAGVSSPIVDDNDSLLDIYCRSDITALAKNRYTHINL 297
QY 380 HQALQLODANTPEFFNGQRCOMCLRSDDLKXVWERLANPGVRRVIVEAGSKRVE 439
DB 298 HQALQLODAYSPEY-IRSORCOMCLRSDDLKXVWERLANPGVRRVIVEAGSKR 356
QY 440 SLSDIEKFL 449
DB 357 SLSDIEKFFI 366

RESULT 10

US-10-424-599-142943
; Sequence 142943, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142943
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(263)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100095C.1.pep
US-10-424-599-142943
Query Match 39.4%; Score 923.5; DB 12; Length 263;
Best Local Similarity 69.9%; Pred. No. 3.8e-83;
Matches 186; Conservative 31; Mismatches 34; Indels 15; Gaps 5;
QY 186 ELETHGSLNLETHETHTISAWKEAKQTNGRNDSQWRP-QOHLVHATPYESLRDIAVKL 244
DB 7 QLGNHGSNLTEELETHETHTISAWKGGK-----WTGFTQCFIRAGPYDNLKELIAVKI 56
QY 245 LONGISTVPIYSSSDGSPQLHLASLGLKICRYFKNSTGNLPILNQPVCISPLG 304
DB 57 LQHGISTVPIH--SEDGSPQLHLASLGLKICRYFNCSSSLPILQIPALPVG 114
QY 305 SWPKIGDLNSRPLMRPNASLSALNMLVOAGVSSPIVDDNDSLLD-TYSRSDITAL 363
DB 115 TWPKIGESNRPLMRPNASLSALNMLVOAGVSSPIVDDNDSLLDIYCRSDITAL 174
QY 364 AKDKVYTHVRLDENTHQAOLQLODANTPEFFNGQRCOMCLRSDDLKXVWERLANPGVR 423
DB 175 AKDRTYTHINLDEMTVHQALQLODQSYNTYE-LSQRCOMCLRTDSLHKVWERLANPGVR 233
QY 424 RVFIVEAGSKRVEGILSLSDIEKFL 449
DB 234 RLIVIVEAGSKRVEGIALSDIEKFFL 259
RESULT 11
US-10-424-599-142947
; Sequence 142947, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142947
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(299)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100095C.1.pep
US-10-424-599-142947
Query Match 39.3%; Score 920; DB 12; Length 299;
Best Local Similarity 62.1%; Pred. No. 1.1e-82;
Matches 187; Conservative 32; Mismatches 36; Indels 46; Gaps 5;
QY 182 LILRELETHGSLNLETHETHTISAWKEAKQTNGRNDSQWRP-QOHLVHATPYESLRDI 240
DB 10 LLKHKLGNHGSNLTEELETHETHTISAWKGGK-----WTGFTQCFIRAGPYDNLKELI 59

RESULT 14
US-10-437-963-125316
; Sequence 125316, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125316
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(166)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27972C.1.pep
US-10-437-963-125316

Query Match 33.2%; Score 778; DB 16; Length 166;
Best Local Similarity 92.1%; Pred. No. 5.9e-69;
Matches 152; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 286 NSTGNLPILNMPVCSYPLGSGWPKIGDLSNRPLAMLRPNASLSALNMLYQAGVSSIPV 345
Db 2 NSQGNLPILSQVCTPLGTWPKIGDPNRPPLAMLRPNATLSALNMLYQAGVSSIPV 61

Qy 346 DENDSLDYSRSDITALAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCQML 405
Db 62 DENDSLDYSRSDITALAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCQML 121

Qy 406 RSPDLKWMERLANPGVRRVFIIVEAGSKRVEGIISLSDIFKLLS 450
Db 122 RSDTLKWMERLANPGVRRVFIIVEAGSKRVEGIISLSDIFKLLS 166

RESULT 15
US-10-425-114-50202
; Sequence 50202, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50202
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700799332_FLI.pep
US-10-425-114-50202

Query Match 32.5%; Score 761; DB 12; Length 477;
Best Local Similarity 38.8%; Pred. No. 1.6e-66;
Matches 173; Conservative 85; Mismatches 152; Indels 38; Gaps 8;

Qy 13 CP-----TVPQACISLSPGHIHYKFFVGDGWRHDEROPTISGEPGIVNTLYLTREYNQIN 67
Db 56 CPMGLVGAFFQVFDLPPGVYQVRFVLDGVYRCDETKPFVRDEYGLISNEVLENNVQPV 115

Qy 68 TLSSPSTPGSRMMVDNENFQRTVTLSDGTV-----SEGLRVSEAAIQISRCRVSEYL 122
Db 116 VOPEPSIRGT--NMD-----KGTILKTMPEPPSSQNPMSQCIATIRHVSGIL 160

Qy 123 NLTCVDLLPDSGVKVIADINLPKQSFHILHEQIPVAPLWDSFRQGFVGLLSPILFIL 182
Db 161 LNTIYDVVPLSSKLTVDLQLPVKQAFKIMHDEGLALVPLWDRQGTITGMLTASDFVL 220

Qy 183 ILRELETHGSLNLTETEQLTHHTISAWKEAKQOTNGRNDQWRPQOHLVHATPYESLRDIAY 242
Db 221 ILRKLQRNIQVIGNEE----PISAWKEAKLQFYGGPOGAAMQRPL:HVKDSDLVDVAL 276

Qy 243 KLLQNGISTVPVIYSSSSDGSFPQLLHLSLGLKICICRYFKNSTGNLPILNQPVCSIP 302
Db 277 TIIRNEISSVP-IFKCMADSSGVPEFLNLTATQGLKFLCSKIQBEAEGCSLHQLLSIP 335

Qy 303 LGSWVPKIGDLSNRPLAMLRPNASLSALNMLYQAGVSSIPVDDNDSLDYSRSDITA 362
Db 336 IGTWSPHTGRSSRQLTLLSSPLNTCLDILLQDRYSSIPVDDNGSLRDVYSLSDIWA 395

Qy 363 LAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCOMCLRSDDLKWMERLANPG 421
Db 396 LAKNDVYARIELEQVTVQNALDVQYQV-----HGRROCHTCLQITLLEVLREGLSIPG 448

Qy 422 VRRVFIIVEAGSKRVEGIISLSDIFKLL 449
Db 449 VRRLVWIEQSTREVEGIISLRDVFLL 476

Search completed: July 7, 2004, 17:53:56
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2004, 05:16:54 ; Search time 5618 Seconds
(without alignments)
3479.478 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREHLPMSPIEGPTVFQAI.....SKRVEGIISLSIDIFKFLSL 451

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh

-Q=/cgn2_1/USPTO.spool/US9857525/runat_07072004_161018_293/app_query.fasta_1.647

-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US9857525@cgn_1_1_3731@runat_07072004_161018_293 -NCPUS=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgc.hum.*
40: em.htgc.mus.*
41: em.htgc.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2315	98.8	2156	8	AF276086	AF276086 Zea mays
2	2218	94.7	2247	8	AF276085	AF276085 Zea mays
3	2085	89.0	2145	8	AK102647	AK102647 Oryza sat
4	2085	89.0	2565	8	AK111501	AK111501 Oryza sat
5	2031	86.7	1632	8	BT009111	BT009111 Triticum
6	1580	67.5	1464	8	BT000625	BT000625 Arabidops
7	1580	67.5	2090	8	AF250335	AF250335 Arabidops
8	1580	67.5	2301	8	AF439826	AF439826 Arabidops
9	1481.5	63.3	1887	8	AY247268	AY247268 Medicago
10	1171.5	50.0	92392	8	AC128647	AC128647 Oryza sat
11	1171.5	50.0	145115	8	AC096690	AC096690 Oryza sat
12	1068	45.6	982	8	BT009623	BT009623 Triticum
13	1003.5	42.8	110514	8	F7G19	AC00106 Sequence
14	798.5	34.1	1923	8	AK121122	AK121122 Oryza sat
15	682	29.1	139089	2	AC146571	AC146571 Medicago
16	661.5	28.2	107857	2	AC146552	AC146552 Medicago
17	380.5	16.2	3067	5	BC060444	BC060444 Xenopus l
18	379.5	16.2	2232	10	AY348865	AY348865 Rattus no
19	377.5	16.1	2192	10	AY348864	AY348864 Mus muscu
20	377.5	16.1	2194	9	AF087875	AF087875 Homo sapi
21	374	16.0	3132	10	BC015283	BC015283 Mus muscu
22	373.5	15.9	1167	9	AB025580	AB025580 Homo sapi
23	373	15.9	987	9	BT007127	BT007127 Homo sapi
24	373	15.9	987	12	BT008166	BT008166 Synthetic
25	373	15.9	1435	6	AR139104	AR139104 Sequence
26	373	15.9	2062	9	HS249976	AJ249976 Homo sapi
27	373	15.9	2203	9	BC020540	BC020540 Homo sapi
28	373	15.9	2223	6	AX877755	AX877755 Sequence
29	373	15.9	2223	6	BD156831	BD156831 Primer fo
30	373	15.9	2223	9	AK001887	AK001887 Homo sapi
31	371	15.8	3427	5	BC059181	BC059181 Dario rer
32	367	15.7	996	9	BT007345	BT007345 Homo sapi
33	367	15.7	1578	6	AR270550	AR270550 Sequence
34	367	15.7	1578	9	HSU42412	U42412 Human 5' -AM
35	367	15.7	1668	9	BC000358	BC000358 Homo sapi
36	366	15.6	3210	3	AY084138	AY084138 Drosophil
37	363.5	15.5	2082	3	AF094764	AF094764 Drosophil
38	363.5	15.5	3157	3	AY166758	AY166758 Drosophil
39	363.5	15.5	3199	3	AY166753	AY166753 Drosophil
40	363.5	15.5	3318	3	AY166757	AY166757 Drosophil
41	363.5	15.5	3679	3	AY166756	AY166756 Drosophil
42	363	15.5	2577	3	BT001510	BT001510 Drosophil
43	363	15.5	2684	3	AY166755	AY166755 Drosophil
44	363	15.5	2687	3	BT003249	BT003249 Drosophil
45	363	15.5	2871	3	AY166754	AY166754 Drosophil

ALIGNMENTS

RESULT 1

```

AF276086      2156 bp mRNA linear PLN 13-AUG-2002
LOCUS        Zea mays protein kinase AKINbetagama-2 mRNA, complete cds.
DEFINITION
ACCESSION   AF276086
VERSION     AF276086.1 GI:11139547
KEYWORDS    Zea mays
SOURCE      Zea mays
ORGANISM    Zea mays
             Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 2156)
            Lumbreiras, V., Alba, M.M., Kleinow, T., Koncz, C. and Pages, M.
            Domain fusion between SNPI-related kinase subunits during plant
            evolution
            EMBO Rep. 2 (1), 55-60 (2001)
JOURNAL     MEDLINE
PUBMED     21146516
PUBMED     11252725
REFERENCE   2 (bases 1 to 2156)
            Lumbreiras, V. and Pages, M.
            Direct Submission
            Submitted (07-JUN-2000) Genetica Molecular, CID,CSIC, Jordi Girona,
            18-26, Barcelona 08034, Spain
FEATURES
             source
             1..2156
             location/Qualifiers
             /organism="Zea mays"
             /mol_type="mRNA"
             /db_xref="taxon:4577"
             /dev_stage="seedling"
             /note="under water stress"
             456..1946
             /note="Snf-4 like"
             /codon_start=1
             /product="protein kinase AKINbetagama-2"
             /protein_id="AA031752.1"
             /db_xref="GI:11139548"
             /translation="MESHGADSAHDAGVCVSSGGATVPFRFVNPYGGKRVFVSGSF"
             RSEHLPMFIEGCPFQAICSLSPGHEIFKFFVGEWRHDERQPTISGEFIVNVL
             YLTREYNQNTLSSPSPGSMNDVNFQRTVLSGDTSGEITLRVSEAAIQISR
             CRVSEYLSHTCTDLPDSKVIALDINLPVKQSFLHBEQGIPEVAPLMDSPRGQVSR
             LLSPLDFILIRLELYHGNLTIEQLETHTSIAWKAQRQTNGRNDQMPQOHLVHA
             TPVEDRDIAVLKNGISTVPVYSSSDGSPOLLHLASLSGLIKICIKRYFKNSTG
             NFPLINQPVCSIFLGSWPKIGLNSRPLAMLRPNASLSALANLWQAGVSSIPVDD
             NSLLDTYSRSDITALAKAYTHYTRIDEMTHIQALQLQGDANFPGFNGQRQCMCL
             RSDPLLKWERLANPGRVRVFIIVEAGSKRVEGIISUSIDIFKFLSL"
CDS
             83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
             837 GTGGATAATGAATTTTCAAGTACGGTTACGTTGTTCAGATGTTACCGTTTCAAGAGT 896
             103 ThrLeuArgValSerGluAlaLalleGlnIleSerArgCysArgValSerGluThrLeu 122
             897 ACTCTGAGAGTTTCAGAGGCTCAATACAAATATCTAGGTGCGCTGTCTGAATATCTG 956
             123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
             957 AGTTTGCATACATGCTATGTTTACTCCAGATTTCTGGCAAGGTATTTGCCCTAGACAT 1016
             143 AsnLeuProValLysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaPro 162
             1017 AATTTACCTGTGAAGCAATCATTTCCATATTTCTCCATGAACAGGGGATTTCTGTAGCTCT 1076
             163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
             1077 CTCTGGGACTCATTCAGAGGTCATTTTGTGGTCTCTTACGCCCATTTGGATTTTCATCTC 1136
             183 IleLeuArgGlnLeuGlnThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHis 202
             1137 ATATTCCGGGAGCTAGAAACTCATGCTCGAATCTGACAGAGAGAGAGCTTGAACACAC 1196
             203 ThrIleSerAlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyr 222
             1197 ACTATATCTGCATGGAAAGAGGCTAAGCGCAACTAATCGAAGAAATGATGGTCAGTGG 1256
             223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGlnSerLeuArgAspIleAlaVal 242
             1257 CGACCGCAACAGCATCTAGTGCATGCCGCCCTTATGAGTCTCTTACGAGGCACTTGCAGTA 1316
             243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
             1317 AAGCTTTTGCAAAATGGCATTTCTACAGTCCAGTATTTATTTATTCATCATCATCATCAGATGGA 1376
             263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
             1377 TCATTTCCCGAGTATTTGCACTTTGCACTCATCTTCTGAAATTTTGAATGATTTGTAGA 1436
             283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
             1437 TACTTCAAAAACCTCAACTGGTAATTTTGCCTTATTTCTGAACCAACCCAGTGTGCTCCATTCCG 1496
             303 LeuGlySerTyrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
             1497 CTGGTTCTCTGGGTTCGGAATAATTTGTTATCTGAACAGTCTGCTCATTTGGCTATGTTGGA 1556
             323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
             1557 CCTAATGCATCACCTTAGCTCTGCCCTTAAACATGTTGTTCAAGCTGGAGTAAGCTCAATA 1616
             343 ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
             1617 CCAATTCTGGATGACACAGCTCTCCCTCTTGCACCTTACTCTAGAGTGACATCACAGCC 1676
             363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
             1677 CTAGCAAAAGCCAGGCTTACACATGTTCCGCTGGATGGATGACCACTTACACAGGCT 1736
             383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGln 402
             1737 TTGCAGCTTGGACAAGATGCCAATACACCTTTTGGATTTCTTTAAGCGCCAGAGATGCCAG 1796
             403 MetCysLeuArgSerAspProLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
             1797 ATGTGCTCTCGGTCTGATCTCTTGTCTGAGGTGATGAGGAGCTGGCTTAATCTCTGGGGTG 1856
             423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer 442
             1857 CGGGGGGTGTTTCATTTGTGAAGCTGGAGGCAACAGCTGTGGAGGCAATCATATCACTAAGT 1916
             443 AspIlePheLysPheLeuLeuSerLeu 451

```

Alignment Scores:

```

Pred. No.:      3 86e-189      Length:      2156
Score:          2315.00      Matches:      446
Percent Similarity: 99.55%      Conservative: 1
Best Local Similarity: 99.33%      Mismatches: 2
Query Match:     98.85%      Indels: 0
DB:              8      Gaps: 0

```

US-09-857-525C-2 (1-451) x AF276086 (1-2156)

```

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
DB 597 GAACATTTGCGGATGTTCTCTATTGAAGGTGCCCACTGTATTTCAGGCTATTTCAGT 656
QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
DB 657 CTGTCTCTGGATTCACAGATGACAAATTTCTTTGGAGCGGGAATGCGGCATGATGAG 716
QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
DB 717 CGTCAACTACCATATCTGGGAGTTTGGCATGTTAACAACATTTACTTACCAAGGAA 776
QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
DB 777 TATAACCAATTAACACCTTCAAGTCCAGACACACCTGGAAGCAGGATGAACATGAT 836

```

```

Db      1917 GATATTTCAAGTTCTTGCTGAGCTTG 1943
|||||
RESULT 2
AF276085      2247 bp mRNA linear PLN 13-AUG-2002
LOCUS      Zea may protein kinase AKINbetagama-1 mRNA, complete cds.
DEFINITION
ACCESSION AF276085
VERSION AF276085.1 GI:11139545
KEYWORDS
SOURCE Zea may
ORGANISM Zea may
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2247)
Lumbreras, V., Alba, M.M., Kleinow, T., Koncz, C. and Pages, M.
Domain fusion between SNF1-related kinase subunits during plant
evolution
EMBO Rep. 2 (1), 55-60 (2001)
MEDLINE 21146516
PUBMED 11252725
REFERENCE 2 (bases 1 to 2247)
AUTHORS Lumbreras, V. and Pages, M.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2000) Genetica Molecular, CID, CSIC, Jordi Girona,
18-26, Barcelona 08034, Spain
FEATURES
source
location/Qualifiers
1..2247
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/dev_stage="seedling"
/notes="under water stress"
496..1989
/notes="Snf-4 like"
/codon_start=1
/product="protein kinase AKINbetagama-1"
/protein_id="AA031751.1"
/db_xref="GI:11139546"
/translation="MFSGADSDAHAGTVGVSGVGATVPARFWFYGGKRVVSGSF
TWSEHPHPSVEGCTVTFQATCSLSPGIHXYFYVDGWRHDERQPTISGEFGIVNT
LYVTRFNFQINALNSTFGSRMNDVNFQHTVTSDDTIPETGTVRSEAAIQIS
RCRVSYLNHTCYDLPSPGKIALDINLPVKQSPHILHEQGPVAPLWDSFRGQV
GLSPDFILILRELTHSGNLTEDOLETHISAWKEAKROTGRNDGOWRAHQLVH
ATPVESLRIDIAVKLLNDISTVPVIYSSSDSGFPOLLHLASGLIKCIFRYEKST
GNLPILNQVCSIPLGSWPKIGCPNRPRLAMLRPNASLSALNMLVQAGVSIPVD
ENDSLDITYSRSDITLAKAKVYTVRLEDEMAIHOALQUGQDANTPFQFNGQRCCMC
JRSDFLLKVMERLANPGRVRFIVEAGSKRVEGIIISLSDIFKFLSL"
CDS
ORIGIN
Alignment Scores:
Pred. No.: 8,95e-181 Length: 2247
Score: 2218.00 Matches: 427
Percent Similarity: 96.88% Conservative: 8
Best Local Similarity: 95.10% Mismatches: 14
Query Match: 94.71% Indels: 0
DB: 8 Gaps: 0
US-09-857-525c-2 (1-451) x AF276085 (1-2247)
QY 3 GluHisLeuProMetSerProLeuGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 640 GAACATTTGCCGATGCTCTGTCGGAAGGCTGCCCCACTGTATTTTCAGGCTATTTCGAGC 699
QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
Db 700 CTGTCTCCAGGGATTCATGAGTCAACAGTTCTATGTGCGAGCGGGAGTGGCTCATGTAG 759
QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLysArgGlu 62
Db 760 CGCCAACCTACTATATCTGGGGAGTTCGGTATAGTTAAACACATATATGACATGACAAAGGAA 819
QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
Db 820 TTTAATCAATAATAATGCTTTATTAATCCAGACACACTGGAAGCAGAGTGAACATGAT 879
QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
Db 880 GTGCATATGAAAATTTTCAACATACGCTTACATTTGTCGGATGTGTACCATTTCCAGAGGT 939
QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
Db 940 ACTGTGAGAGTTTCAGAGGCTGCAATACAAATCTCAAGGTGCCGCTTCTGAGTATCTC 999
QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 1000 AATTGTGATACATGCTATGATTACTCCGGATTCCGGCAAGGTTATGTCTCTGGACATT 1059
QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db 1060 AATTACTCTGGAAGCAATCTTTTCATATCTGCATGAACAGGGGATTCCTGTAGTCTCT 1119
QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 1120 CTCTGGGACTCATTCAGAGGTCAAATTTCTCGGCCCTTCTGAGCCCATTTGGATTTTCATCTT 1179
QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHis 202
Db 1180 ATATTGGGGAGCTAGAACTCATGGCTCGNACTTGACAGAGATCAGCTTGAAACACAC 1239
QY 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
Db 1240 ACTATATCTGCATGGAAGAGGCTTAGCGGCAAACTTGTGGAAGAAATGATGTCAGTGG 1299
QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
Db 1300 CGAGCACATCAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGAGCATTCAGATA 1359
QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
Db 1360 AAGCTTTTGTAAATGACATTTCTACAGTCCAGTTATTTATTCATCATCATCAGATGGA 1419
QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
Db 1420 TCATTCCTCAGTTATTCAGCTTCCTGATCCCTTTCTGSAATTTTGAATGATTTATTTAGG 1479
QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
Db 1480 TATTTTAAAACTCAACTGGTAAATTTGCCATTATCTGAACCAACCGGCTGCTCATTTCCG 1539
QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
Db 1540 CTGGTTCTTGGGTTCCGAAATCCGGTGATCCAAACAGTCTCCATTGGCTATGTCGA 1599
QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
Db 1600 CCTAATGTCATCACTTAGCTCTGCCCTTAACATGTGGTTCGATGAGTGGAGTCACTAATA 1659
QY 343 ProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
Db 1660 CCAATGTGGATGAAACGACTCCCTGCTTGACACTTACTCTAGAGTGAATCATCAGACC 1719
QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
Db 1720 CTAGCTAAAGCCAAAGCTACACACATGTTCCGCTGGATGAGATGGCCATTTCATAGGCT 1779
QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
Db 1780 TTACAGCTTTGACACAGATGCCAATACGCTTTTGGATTTTAAACGCTCAGAGATGCCAG 1839
QY 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
Db 1840 AUGTGCTTCGGTCTTGATCTCTTGTGAGAGTGAATGAGGAGATGGCTTAATTCGGGGGTG 1899
QY 423 ArgArgValPheIleValGlnAlaGlySerLysArgValGluGlyIleIleLeuSerLeuSer 442

```


QY	143	AsnLeuProValLysGlnSerPheHisIleLeuHisGlnGlnGlnProValAlaPro	162
Db	624	AAITTCCTGTGAAGCAATCTTTTCATATCTTCATGAACAGGGAATCTCTGTGGCACT	683
QY	163	LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu	182
Db	684	CTATGGGATTTCATCGAGGGCCAGTTTGTGGCTTTTGAGCCCACTGGATTTTATATC	743
QY	183	IleLeuArgGlnLeuGlnThrHisGlySerAsnLeuThrGlnGlnLeuGlnThrHis	202
Db	744	ATATTGAGAGAGCTGGAAATCTAGCTCCAAATCTGACAGAGAGCACTGAAACACAT	803
QY	203	ThrIleSerAlaTrpLysGlnAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp	222
Db	804	ACTATATCTGCGTGGAGGAGGCAAGCTCAAACTTATCCAGAAAGAGGGTCTTGG	863
QY	223	ArgProGlnGlnHisLeuValHisAlaThrProTyrGlnSerLeuArgAspIleAlaVal	242
Db	864	AGGGCAAAATCACCATTAGTTGATGTCACCCCTTATGAATCCCTGAGGAAATTCCTATG	923
QY	243	LysLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly	262
Db	924	AGATACTCGAAACGGTGTCTACCGTTCCAAATTAATGTTTCTCTCATCACCAGATGGC	983
QY	263	SerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg	282
Db	984	TCATATCCCAATGTTCATCTTGCCTTCCTTCTGGAATTTGAATGTAATTTGTAGA	1043
QY	283	TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro	302
Db	1044	TATTTTAAATAATCTCAAGGTAATTTTACCTATTTTGAGCAACCTGTATGATCAAAATTCCT	1103
QY	303	LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg	322
Db	1104	CTGGTACCTGGGTTCAAATAATGGTGATCCCTAATGGCGCTCATGGGTATGTTCCGG	1163
QY	323	ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle	342
Db	1164	CCTAACACATCTCTTAGCGTGGCTGCAATTTCTGCTTCAAGCTGGTGAGCTCAATA	1223
QY	343	ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla	362
Db	1224	CCAAITTTGGATGACAAATGACTCACTGCTTGACACATATTTCCAGAGTGCATCACAGCT	1283
QY	363	LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla	382
Db	1284	CTTGCNAAAGCAAGGTGTACACACATCTCGCTAGATGAGATGACCATTCATCAGGCT	1343
QY	383	LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGln	402
Db	1344	CTGCAGCTTGACAGGATGCAATTCGCCCTTTGGATTTCTTAACGGGCAAGATGTCAG	1403
QY	403	MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal	422
Db	1404	ATGTGTCCTGGTCTGACATTTGTGAAAGTAGTGAGGCGATTAGCTAATCTCTGGGGTG	1463
QY	423	ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer	442
Db	1464	CGCCGTTTTCATTTGTGGAGCTGGTAGCAACCGTGTGGAGGCGATAATATCATCTAGT	1523
QY	443	AspIlePheLysPheLeuLeuSer 450	
Db	1524	GATATTTTCAAGTTCTTGCTGAGC 1547	
RESULT 4			
LOCUS	AK111501		
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J013000N04, full insert sequence.		
ACCESSION	AK111501		
VERSION	AK111501.1 GI:37988164		
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		

ORGANISM

Oryza sativa (japonica cultivar-group)
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Cryzeae; Oryza.

REFERENCE

AUTHORS

1
The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team;
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, M., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Ohtsuki, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

JOURNAL

MEDLINE

PUBMED

AUTHORS

2
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, Y., Ikeda, R., Inamura, K.,
Iotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahara, Y.,
Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K.,
Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M.,
Oka, H., Osato, N., Ota, Y., Ohtsuki, R., Ryu, R., Saitoh, H., Sakai, C.,
Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y.,
Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tadami, M.,
Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K.,
Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,
Yokomizo, S. and Yoshimura, A.

Rice full-length cDNA

Unpublished

3 (bases 1 to 2565)

Kikuchi, S.

Direct Submission

Submitted (12-SEP-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
tsi:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,
Yamamoto, M. and Nakahama, Y.
PAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M.,
Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M.,
Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M.,
Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,
Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

TITLE

JOURNAL

AUTHORS

JOURNAL

COMMENT

FEATURES

Location/Qualifiers
 1..1632
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="w11.pk0008.b3.fis"

ORIGIN

Alignment Scores:
 Pred. No.: 6,84e-165 Length: 1632
 Score: 2031.00 Matches: 386
 Percent Similarity: 93.74% Conservative: 33
 Best Local Similarity: 86.35% Mismatches: 26
 Query Match: 86.72% Indels: 2
 DB: 8 Gaps: 1

US-09-857-525C-2 (1-451) x BT009111 (1-1632)

QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
 DB 2 CACGACCCGATGTCCTCGAGTCGAGGTTGCCCACTGTATTTCAGCTATCTGCACTTA 61
 QY 24 SerProGlyIleHisGluTyrIysPheValAspGlyGluTrpArgHisAspGluArg 43
 DB 62 CCTCCAGGCGATTATCAGTACAAAGTTCAATGTGCACGCGGCGAGTGGAGGCATGACGAGGG 121
 QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyr 63
 DB 122 CAACCTACTATAACTCGAGAGTATGGGTGGTAAACACTTATATCTGACAAAGGGAATTT 181
 QY 64 AsnGlnIleAsnThrLeuSerProSerThrProGlySerArgMetAsnVal 83
 DB 182 GACCACATTAATCTGTACGACCCCACTACCTCGGAGCAGG-----ATGGATGTG 235
 QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
 DB 236 GACAGTGACAGTATTCAACGAATGGGTTCGTTCGGATGGTGCCTTCAGGAAGGTCTCT 295
 QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsn 123
 DB 296 CCAAGAATCTCAGAGGCTGTCTACAGATCTTAGGTGTCTGTGTCTGAGTATCTGAAAT 355
 QY 124 LeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
 DB 356 GCGCATACAGGCTATGACCTACTACAGATCTTGGAAAGGTCATTGCTCGGCATTAAT 415
 QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyLeuProValAlaProLeu 163
 DB 416 TTACCTGTGAAGCAATCTTCCATATTCCTCATGAACAGAGGGATTCCTGTGGCTCCTCTG 475
 QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
 DB 476 TGGGATTCATTCAGGGGTCAAGTTGTTGGCTTCCTGAGCCACCTGGATTTTATCTATTA 535
 QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThr 203
 DB 536 TTCAGAGAGCTGGAACCTCATGGCTCAAACTGACAGAGAAACAGCTTGAACACACACT 595
 QY 204 IleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArg 223
 DB 596 ATATCTCGCTGGAAGAGGCTAAGCGGCAAACTTATGGAAGAATGATGGACAACTTAGA 655
 QY 224 ProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLys 243
 DB 656 TCAAAATCAGCATCTAGTGTGATGCCACCCCTTATGAAATCCTTACGGGGTATTCGCATGAA 715
 QY 244 LeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGlySer 263
 DB 716 ATACTCGAAACTGCACTTCTACAGTCCCAATCATCTATTTCATCATCAGATGGATCG 775
 QY 264 PheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyr 283
 DB 776 TTTCGGCAGCTGTGTGATCTTGTGATCCCTTTTCAAGGAATTTTGAATGATCTGTAGATAC 835

QY 284 PheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeu 303
 DB 836 TTCAGAACTCCACTGGTGTGTTGCGGATTTCTAAACCAACCACTATGCTCAATTCGCTG 895
 QY 304 GlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgPro 323
 DB 896 GGTACCTGGGTTCCAAAAATTTGGTGAACCAATAGGTTCATCCATTTGGCTATGTTGCGGCT 955
 QY 324 AsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIlePro 343
 DB 956 AATACATCTCTTAGCTTCGCCCTTAACCTTGTGTTCAAGCTGAGTAGTTCATACCC 1015
 QY 344 IleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeu 363
 DB 1016 ATTGTGGATGATTAACGACTCGCTGATCGACATATCTCCAGAGTGCATCACAGCTCTA 1075
 QY 364 AlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeu 383
 DB 1076 GCGAAAGACAAAGGCTACACCCATATCCGCTAGATGAGATGACCATTCATCAGGCTTG 1135
 QY 384 GlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheSerGlyGlnArgCysGlnMet 403
 DB 1136 CAGCTCGGCGACAGCCGAATTCACCTTTTGGACTTTTCAATGGTCAAGAGATGCCAGATG 1195
 QY 404 CysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArg 423
 DB 1196 TGCTCCAGTCTGACCTTTGCTGAAGTTATGAGAGATTGGCTAATCTCTGGGCTGCT 1255
 QY 424 ArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAsp 443
 DB 1256 CGCGTGTTCATCTGAGGCTGGCAGCAGCGAGTGGAAAGCGTAAATATCTGCTGAGCGAC 1315
 QY 444 IlePheLysPheLeuLeuSer 450
 DB 1316 ATATTCAAGTTGCTGCTGAGC 1336

RESULT 6

BT000625

LOCUS 1464 bp mRNA linear PLN 25-SEP-2002
 DEFINITION Arabidopsis thaliana putative activator subunit of SNF1-related protein kinase SNF4 (At1g09020/F7G19_11) mRNA, complete cds.

ACCESSION

BT000625.1 GI:23308442

VERSION

BT000625

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 1464)

AUTHORS

Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE

Arabidopsis ORF clones

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1464)

AUTHORS

Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE

Direct Submission

JOURNAL

Submitted (25-SEP-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN Arabidopsis Full-length cDNA"): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PSEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Shinn, P., Banu, J., Bowser, L., Chan, M. M., Chang, E., Dale, J. M., Deng, J. M., Goldsmith, A. D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers
 1. .1464
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="U14970"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51"
 1. .1464
 /note="putative activator subunit of SNF1-related protein kinase SNF4"
 /codon_start=1
 /product="Atg39020/F7G19_11"
 /protein_id="AA18191.1"
 /db_xref="GI:23308443"
 /translation="MFQSTLDSRRNSAASGQLLPTFRFVMPYGGRRVFLSGSPTRWT EHYPMSPLESCPTPQVQCNLTGCHYOYKFPDGEWHDHDPVPSGNGGVNTIFIT GPVMPAGSPFELGRNMDVDVFLRTADPSQBAVERMSGVLDLESRRLSVLLSFR TAYVLEPSKQALNDVLPVKQAFHLYEQGIPFLAPLDFPKGQFVGVLPLOFLI LRELTHGSLNTEELVEYSLQDGYKQAHISRDYSGRPYRPLPVQVGVNDLKV ALKILQNAAPVYISLQDGYKQAHISRDYSGRPYRPLPVQVGVNDLKVQPI CSPPLGTVPRIGSSKPLMTPLPHASLGSLALYQAEVSSI PVVDNDNDLSDIYS RSDITLAKQKAYAKHLDDMTVHQALQGDQSPFVGFNGRQCHMCLRSLSLVKVM ERLANPVRRLVIVEAGSKRVEGIISLSDVPEQLLS."

CDS

1. .1464
 /note="putative activator subunit of SNF1-related protein kinase SNF4"
 /codon_start=1
 /product="Atg39020/F7G19_11"
 /protein_id="AA18191.1"
 /db_xref="GI:23308443"
 /translation="MFQSTLDSRRNSAASGQLLPTFRFVMPYGGRRVFLSGSPTRWT EHYPMSPLESCPTPQVQCNLTGCHYOYKFPDGEWHDHDPVPSGNGGVNTIFIT GPVMPAGSPFELGRNMDVDVFLRTADPSQBAVERMSGVLDLESRRLSVLLSFR TAYVLEPSKQALNDVLPVKQAFHLYEQGIPFLAPLDFPKGQFVGVLPLOFLI LRELTHGSLNTEELVEYSLQDGYKQAHISRDYSGRPYRPLPVQVGVNDLKV ALKILQNAAPVYISLQDGYKQAHISRDYSGRPYRPLPVQVGVNDLKVQPI CSPPLGTVPRIGSSKPLMTPLPHASLGSLALYQAEVSSI PVVDNDNDLSDIYS RSDITLAKQKAYAKHLDDMTVHQALQGDQSPFVGFNGRQCHMCLRSLSLVKVM ERLANPVRRLVIVEAGSKRVEGIISLSDVPEQLLS."

ORIGIN

Alignment Scores:
 Pred. No.: 3,41e-126 Length: 1464
 Score: 1580.00 Matches: 304
 Percent Similarity: 81.78% Conservative: 64
 Best Local Similarity: 67.56% Mismatches: 74
 Query Match: 67.46% Indels: 8
 DB: 8 Gaps: 5
 US-09-857-525C-2 (1-451) x BT000625 (1-1464)
 QY 3 GluHisLeuProMetSerProIleGlyCysProThrValPheGlnAlaIleCysSer 22
 DB 133 GAGCATGTGCGCAATGTACCGCTCGAGGCTGCCCTACTGTCTTTTCAAGTCATTGCAAC 192
 QY 23 LeuSerProGlyIleHisGlyTyrIlePheValAspGlyGluTyrArgHisAspGlu 42
 DB 193 TTGAGCCAGCATATCATCAGTATAAGTTTGTGTGATGGGAATGCGGCACAGTACG 252
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuArgGlu 62
 DB 253 CACCAACCATTTGTAAGCGGAATGCTGGAGTAGTGAATACATAATATTATACTGGGCCA 312
 QY 63 TyrhanGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 DB 313 ---GATATGTTCTGCTGTTTATAGCCCGAGACACTGGGC---CGGTGCAATATGGAT 366

QY 83 ValAspAsnGluAsnPheGlnAcqThrValThrLeuSerAspGlyThrValSerGluGly 102
 DB 367 GTGGAT---GATGTCCTTCGCGAAGCGCTGACCCGCTCC-----CAGGAAGCT 411
 QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 DB 412 GTTCTAGGATGTCGGGGGTGATTGCGAGCTTCTCGTCACCGCATATCGTTTATTG 471
 QY 123 AsnLeuHisThrCysTyrAspLeuProAspSerGlyIleValIleAlaLeuAspIle 142
 DB 472 TCACCCGCGCATGTCATATGAGCTCTCCCAAGTCGGGCAAGGTATTGCAATGGATGT 531
 QY 143 AsnLeuProValIleGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 DB 532 AATTTCAGTGAAGCAGCATTCATATATCTATGAGCAGGGAATCCCTTTGGCTCT 591
 QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 DB 592 CTTTGGGACTTTCGAAAGGCGCAATTGTTGGAGTCTCTGGTCCACTAGACTTCATTCTA 651
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuThrHis 202
 DB 652 ATACTGAGAGAGCTTGGAACTCATGGATCAAACTTGACAGAGGAGAGCTTGAGACGGAC 711
 QY 203 ThrIleSerAlaTrpIleGluAlaIleArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
 DB 712 ACATAGCAGCTCGAAAGAGGGGAGGCTCATATTAGCCGCAATATGATGGAAGTGG 771
 QY 223 ArgPro---GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAla 241
 DB 772 AGACCATATCTAGGCCACTTGTTCAGGTTGGAGCCCTATGATATCTGAAAGAGCTGCC 831
 QY 242 ValIleLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSer 261
 DB 832 CTGAAATTTTGCAGAACAGAGTGGCTGCTTCGCTATCAGGCATATTAATGATATG 891
 QY 262 GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
 DB 892 GGTTTCATATCCGAGTTACTTCATCTTCTTCGCTATCAGGCATATTAATGATATG 951
 QY 282 ArgTyrPheIleAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
 DB 952 AGATACCTTTAGACATCTGCTAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1011
 QY 302 ProLeuGlySerTrpValProIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321
 DB 1012 CCCTCGGTAGTGGTCCCTAGATCGAGATCAAGTAGCAAACTCTCGTACATTG 1071
 QY 322 ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
 DB 1072 AGACTCAGGCTCTTTGGGTTCGCGCTCGCATTTATTAGTTCAAGCTGGAAGTCAGTTC 1131
 QY 342 IleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr 361
 DB 1132 ATTCCGTAGTGGATGACACGACTCGCTTATTGACATATCTCTCGAAGTGAATTA 1191
 QY 362 AlaLeuAlaIleAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGln 381
 DB 1192 GCTCTGGCTAAGACAGGCGCATACGACAGATTCTTGTATGATGACATGACAGTTCAC 1251
 QY 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
 DB 1252 CGCTACAGTTGGGCAAGATCGAGCGCGCTTATGGAATCTTCAACGGGCGAGATGT 1311
 QY 402 GlnMetCysLeuArgSerAspProIleLeuValMetGluArgLeuAlaAsnProGly 421
 DB 1312 CACATGTCTTGGCTCAGACTCTCTGTGAAGATGATGGAGCGGTTCGCGAATCCAGT 1371
 QY 422 ValArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeu 441
 DB 1372 GTAAGGAGGCTGTGTATGATGGAAGCGGAGCAACCTGTTGAAGTATCATTTCACTG 1431
 QY 442 SerAspIlePheLysPheLeuLeuSerLeu 451

|||||:::|||||:::|||||:::|||||

Db 1432 AGCGATGTTTCCAAATTCTGCTGGGTCTT 1461

RESULT 7
AF250335 LOCUS
DEFINITION
Arabidopsis thaliana 2090 bp mRNA linear PLN 06-SEP-2000
protein kinase SNF4 mRNA, complete cds.
AF250335
AF250335.1 GI:9965728
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 2090)
Kleinow,T., Bhalarao,R., Breuer,F., Umeda,M., Salchert,K. and
Koncz,C.
Functional identification of an Arabidopsis snf4 ortholog by
screening for heterologous multicopy suppressors of snf4 deficiency
in yeast
Plant J. 23 (1), 115-122 (2000)
20387008
10929106
JOURNAL MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 2090)
Kleinow,T., Bhalarao,R. and Koncz,C.
Direct Submission
Submitted (29-MAR-2000) Genetic Principles of Plant Breeding,
Max-Planck-Institute for Plant Breeding, Carl-von-Linne-Weg 10,
Koeln 50829, Germany
LOCATION/Qualifiers
1. .2090
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/map="1-8..7"
932..2080
CDS
/note="contains 4 cystathione-beta-synthase (CBS) domains;
heterologous multicopy suppressor of snf4 deficiency in
yeast"
/codon_start=1
/product="putative activator subunit of SNF1-related
protein kinase SNF4"
/protein_id="AAGI0141.1"
/db_xref="GI:9965729"
/translations="MVPAGFSPTLGRSNMDVDVRLTRADPSQEAVERPMGVLDLELS
RHRISLLSTRAYELLPESGLIDLVNLPVKQAFLHYBOGIPAPLWDFKGQFY
GVLPGLDILIMRELGHSTTEBELETHIAAWKEKAHISRQDGSRPYPLRV
QVGPNLQVALIKONKVAAPVIVYSLQDGSYPQLLAHLASLGILKICRYPRHS
SSSLPILOPICSPIGTWPRIGESSKPLATLRPHASLGALALLVQAEYSIPWV
DNDSLIDIYRSDDITALAKDAKAYAHILDMDTVHQALQGDASPYPYGFNGQRCHM
CURSDSLVXMERLANPGVRRLVIVENGSKVEBIIISLSDVFQFLGL"

ORIGIN
Alignment Scores:
Pred. No.: 5,43e-126 Length: 2090
Score: 1580.00 Matches: 304
Percent Similarity: 81.78% Conservative: 64
Best Local Similarity: 67.56% Mismatches: 74
Query Match: 67.46% Indels: 8
DB: Gaps: 5

US-09-857-525C-2 (1-451) x AF250335 (1-2090)

Qy 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
|||:::|||||:::|||||:::|||||
Db 749 GAGCAUTGCCAATGTCACCGCTCGAGGGTGCCCTACTGTTTCAAGTCANTTGCAC 808
|||:::|||||:::|||||:::|||||

Qy 23 LeuSerProGlyIleHisGluTyrlYspPheValAspGlyGluTrpArgHisAspGlu 42
|||:::|||||:::|||||:::|||||

Db 809 TTCACGCCAGCATATCATCATGATAAGTTTTTTTGTGATGGGAATGGCGCACGATGAG 868
|||:::|||||:::|||||:::|||||

Qy 43 ArgGlnProThrIleSerGlyCluPhedGlylleValasnThrLeuTyrlLeuThrArgGlu 62
|||:::|||||:::|||||:::|||||

Db 869 CACCACCAATTTGTAAAGCGGAATGGTGGAGTAGTAACAATAATTTATACTGGGCCA 928
|||:::|||||:::|||||:::|||||

Qy 63 TyrAsnGlnIleAsnThrLeuSerProSerThrProGlySerArgMetAsnMetAsp 82
|||:::|||||:::|||||:::|||||

Db 929 ---GATATGGTCTCTGCTGGTTTTAGCCCGAGACACTGGGC---GGTTCGATAATGAT 982
|||:::|||||:::|||||:::|||||

Qy 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
|||:::|||||:::|||||:::|||||

Db 983 GTGGAT---GATGTCTTCTGCGAACGGCTGACCCGCTC-----CAGGAAGCT 1027
|||:::|||||:::|||||:::|||||

Qy 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrlLeu 122
|||:::|||||:::|||||:::|||||

Db 1028 GTTCTCTAGGATGTCGGGGGTTCATTGGAGCTTCTCGTCACCGCATATCGGTTTTATTG 1087
|||:::|||||:::|||||:::|||||

Qy 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
|||:::|||||:::|||||:::|||||

Db 1088 TCACCCGCACTGCATHTAGCTGCTCCAGAAATCGGCAAGGTATTTCATTTGGATGTT 1147
|||:::|||||:::|||||:::|||||

Qy 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
|||:::|||||:::|||||:::|||||

Db 1148 AATTGGCAGTGAAGCAAGCATTTCCATATCTATGAGCAGGAATCCCTTTGGCTCCT 1207
|||:::|||||:::|||||:::|||||

Qy 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
|||:::|||||:::|||||:::|||||

Db 1208 CTTTGGGACATTTGGAAAAGGCCAAATTGTGGAGTTCTTGGTCCACATGATTCATCTCA 1267
|||:::|||||:::|||||:::|||||

Qy 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHis 202
|||:::|||||:::|||||:::|||||

Db 1268 ATACTGAGAGAGCTTGGAACTCATGGATCAAACCTTACAGAGAGAGAGCTTGAGAGCAC 1327
|||:::|||||:::|||||:::|||||

Qy 203 ThrIleSerAlaTriplysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
|||:::|||||:::|||||:::|||||

Db 1328 ACATAGCAGCCTGGAAAGAGGGGAGGCTCATATTAGCGCGCAATATGATGGAAGTGG 1387
|||:::|||||:::|||||:::|||||

Qy 223 ArgPro---GlnGlnHisLeuValHisAlaThrProTyrlGluSerLeuArgAspIleAla 241
|||:::|||||:::|||||:::|||||

Db 1388 AGACCATATCTTAGCGCACTTGTTCAGGTGGACCCCTATGATAAATCTGAAGAAGCTTGC 1447
|||:::|||||:::|||||:::|||||

Qy 242 ValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrlSerSerSerSer 261
|||:::|||||:::|||||:::|||||

Db 1448 CTCAAAAATTTGCAAAACAAAGGTGGCAGCTGTTCGGTTATATATCTCTTTTGAGGAT 2507
|||:::|||||:::|||||:::|||||

Qy 262 GlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
|||:::|||||:::|||||:::|||||

Db 1508 GGTTTATATCCGAGTTACTGCACTTGTCTGCTATCAGGCATATATAAATGATATGTC 1567
|||:::|||||:::|||||:::|||||

Qy 282 ArgTyrlPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
|||:::|||||:::|||||:::|||||

Db 1568 AGATACTTTAGACATTCGTCTAGCTCTTACCTATCTCTTCAAGCAACCCATTTGTTCAAT 1627
|||:::|||||:::|||||:::|||||

Qy 302 ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321
|||:::|||||:::|||||:::|||||

Db 1628 CCCCTGGTACGTGGGTCTCTAGATCGGAGAAATCAAGTAGCAAACTCTTGTCTACATG 1687
|||:::|||||:::|||||:::|||||

Qy 322 ArgProAsnAlaSerLeuSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
|||:::|||||:::|||||:::|||||

Db 1688 AGACCTCAGCCCTCTTTGGGTCTGGGCTCGCATATTATGTTCAAGCTGAGTCAGTTCA 1747
|||:::|||||:::|||||:::|||||

Qy 342 IleProIleValAspAspAsnSerLeuLeuAspThrTyrlSerArgSerAspIleThr 361
|||:::|||||:::|||||:::|||||

Db 1748 ATTCCCGTAGTGGATGACAACGACTCGCTTATTGACATATATCTCTCGAAGTGATAA 1807
|||:::|||||:::|||||:::|||||

Qy 362 AlaLeuAlaLysAspLysValTyrlThrHisValArgLeuAspGluMetThrIleHisGln 381
|||:::|||||:::|||||:::|||||

Db 1808 GCTCTGGTAAAGACAAGGCATACGACACAGATTTCATTTGATGACATGACAGTCCACAG 1867
|||:::|||||:::|||||:::|||||

Qy 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
|||:::|||||:::|||||:::|||||

Db 1868 GGCTACAGTTGGGCAAGATCGACCCGCTTATGGAAATCTTCAACGGCGAGAGATGT 1927
|||:::|||||:::|||||:::|||||

```

/clone="RAFL09-86-B11 (R14970)"
/ecotype="Columbia"
1. .603
604. .2067
/note="putative activator subunit of SNF1-related protein kinase SNF4"
/codon_start=1
/product="Atlg90920/F7G19_11"
/protein_id="AAL27498.1"
/db_xref="GI:16612254"
/translation="MFGSLDSSRGNSSAASGQLLTFTRFVWYGGRRVLSGSFTRWT
EHPMSPLGECPTVFQVILNLTGPHQYKFFVDGWRHDEHQFVFSNGGVVNTLIT
GPDMPAGSPETLGRSNDVDVFLRTADSQEAPVRMSGVDSLRLSVLSLST
TAYELLPSGKVIALDNLVPKQAFHILVEQGIPLAPLWDFGQGFVGLVGLDILL
LRELTGHSNLTPEELETHITAAWKEGKAHISROYDGRCPYPRPLVOVQPYDNLKDV
ALUKIQNKVAPVPIYSSLDGSSYQPLHLSLGLKLCRYFRHSSSLPLIOPI
CSPIGLTWPVRIIGESSKPLATLRPHSLGSLALLVQAEVSSIIPVDNDLSLIDYS
RSDITALAKRAYAKHLDDMTVHQALQGDASPPYGIPIFGQRCHMCLRSLSLVKM
ERLANPVRRLVIVEAGSKRVEGIIISLSVDFQFLGL"
misc_difference 1267
/note="compared to genome"
/replaces="g"
2068. .2301
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 6,17e-126 Length: 2301
Score: 1580.00 Matches: 304
Percent Similarity: 81.78% Conservative: 64
Best Local Similarity: 67.56% Mismatches: 74
Query Match: 67.46% Indels: 8
DB: Gaps: 5
US-09-857-525C-2 (1-451) x AF439826 (1-2301)
QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 736 GAGCATGTGCCAATGATCAGTCCGCTCGAGGGCTGCCCTACCTGTTTTCAGTCATTCGCAAC 795
QY 23 LeuSerProGlyIleHisGluTyrLysPheValAspGlyGluTrpArgHisAspGlu 42
Db 796 TTGACGCCAGGTATCATCAGTATAAGTTTTTTTGTGATGGCAATGCGCGCACGATGAG 855
QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
Db 856 CACCAACCATTTGTAAGCGGAATGTTGAGTAGTGAATACATATTTATTAATCTGGGCCA 915
QY 63 TyrAsnGlnIleAsnThrLeuSerProSerThrProGlySerArgMetAsnMetAsp 82
Db 916 ---GATATGTTCTCTGCTGTTTATGCCACAGACACTGGGC---CGTCCGAATATGGAT 969
QY 83 ValAspAsnGlnAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
Db 970 GTGGAT---GATGTCTCTCCGCAACGGCTGACCCCGTCC-----CAGGAGCT 1014
QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
Db 1015 GTTCTAGGATGTGGGGGTGATTTGGAGCTTTCGTCCACGCATATCGTTTTATTG 1074
QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 1075 TCAACCGCACTGCATATGAGCTGCTCCAGAAATCGGGCAAGGTTATTCATTTGGAGTGT 1134
QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db 1135 AATTGCGCATGAGCAAGCATTCATATCTATGAGCAGGGAATCCCTTTGGCTCT 1194
QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 1195 CTTTGGGACTTTGGAAAAGGCCAATTTGTTGGAGTTCTTGTGCTCCACTAGACTTCATCTA 1254
QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis 202

```


Db	1255	ATACTGAGAGAGCTTGSAACTCATGATCAAACTTCACAGAGGAGAGCTTGAGAGGCAC	1314	Unpublished	
Qy	203	ThrIleSerAlaTrpIysGluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrp	222	2 (bases 1 to 1887)	
Db	1315	ACAATAGCAGCTTGAAGAGGGAAGGCTCATATTAGCGGCAATATGATGGAATGGG	1374	Buitink, J. and Leprince, O.	
Qy	223	ArgPro---GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAla	241	Direct Submission	
Db	1375	AGACCATATCTTAGGCCACTTGTTGAGGTGGACCTATGATAAATCTGAAGAGCTGGC	1434	Submitted (03-MAR-2003) UMR 1191 Molecular Seed Physiology, INRA,	
Qy	242	ValIysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAsp	261	16 Bd Lavoisier, Angers 49045, France	
Db	1435	CTGAAATTTTGCAAAACAGGTGGCAGCTGTTCCGTTATATATCTTCTTTGCAGGAT	1494	Location/Qualifiers	
Qy	262	GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuIysCysIleCys	281	1. .1887	
Db	1495	GGTTTCATATCCGAGTTACTGTCATCTTGCTTCGCTATCAGGCATATTAAATGATATGC	1554	/organism="Medicago truncatula"	
Qy	282	ArgTyrPheIysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle	301	/mol_type="mRNA"	
Db	1555	AGATACCTTTAGACATTCGCTAGCTCTTTACCTATCTTCAGCAACCCATTTGTTCAAT	1614	/cultivar="Paraggio"	
Qy	302	ProLeuGlySerTrpValProIysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu	321	/db_xref="taxon:3880"	
Db	1615	CCCTGGGTACGTGGTCCCTAGATCGGAGATCAAGTAGCAAACTCTCGCTACATTG	1674	/dev_stage="embryo"	
Qy	322	ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer	341	65. .1522	
Db	1675	AGACCTCAGCCCTCTTTGGGTTCTCGCTCGCATATTAGTTCAAGTGAAGTCAGTTCA	1734	/codon_start=1	
Qy	342	IleProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr	361	/product="AKIN betagamma"	
Db	1735	ATTCCCGTAGTGGATGACACAGCTCGCTTATTGACATATCTCTCGAATGATATAACT	1794	/protein_id="AA061673.1"	
Qy	362	AlaLeuAlaIysAspIysValTyrThrHisValArgLeuAspGluMetThrIleHisGln	381	/db_xref="GI:29887975"	
Db	1795	GCTCTGGCTAAAGCAAGGCATACGCACAGATTCTCTTGATGACATGACAGTTCAACCAG	1854	/translation="MYPSLTDVGGSSSSLSSSSSGPIIPKRFVWPVGCTRVYLGISF	
Qy	382	AlaLeuGlnLeuGlyGnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys	401	TRMSEHIPMSMEGCPVFQVCSLMPGHQFKFNVDQWRIDEQDFVNGVNGIYNT	
Db	1855	GCCTACAGTTGGGGCAAGATGCGAGCCCGCTTATGGAATCTTCAACGGGCGAGATGT	1914	IYIVRRPDILPALISATSSSRSHMEVDNDFGHAEANPRMSDELSRRRIKXFLSE	
Qy	402	GlnMetCysLeuArgSerAspProLeuLeuIysValMetGluArgLeuAlaAsnProGly	421	HTAYDLPSGKVIALDVALPVKQAEHLVYEQDVSMAPLDFCKSQVGVLSAMDPLK	
Db	1915	CACATGTGCTTGGCTCAGACTCTCTTGTAAGATGATGGAGCGGTGGCGAATCCAGGT	1974	ILKELGTGSHLNEQLETHITIAWKEGSKORRALDNNSGPHCFVHAGPKELKD	
Qy	422	ValArgArgValPheIleValGluAlaGlySerIysArgValGluGlyIleIleSerLeu	441	VALKLVQNKVSTVPIISLEDGSPQLLRLPNAISGDALSMFYQAKVSSPIVDENSLDLOYSR	
Db	1975	GTAAAGAGGCTGGTGTAGTGAAGCAGGAGCAACGTTGTGAAGGTATCATTTCACTG	2034	SIPILGTWVNDGPNQGLIRLPNNAISGDALSMFYQAKVSSPIVDENSLDLOYSR	
Qy	442	SerAspIlePheIysPheLeuSerLeu	451	SDITLAKDKAVARIISLDETNIHQALILGQDANSFYGLNNHGRCHMCLRSLSLHKWE	
Db	2035	AGCGATGTTTTCATTTCTGCTCGGTCCT	2064	RLAKPGVRRLVIVEAGSKRVEGIIISLSDVFRFLIG"	
RESULT 9					
AY247268		1887 bp	mRNA	linear	PLN 15-APR-2003
LOCUS	Medicago truncatula				
DEFINITION	Medicago truncatula AKIN betagamma mRNA, complete cds.				
ACCESSION	AY247268				
VERSION	AY247268.1	GI:29887974			
KEYWORDS	Medicago truncatula (barrel medic)				
SOURCE	Medicago truncatula				
ORGANISM	Medicago truncatula				
REFERENCE	1 (bases 1 to 1887)				
AUTHORS	Buitink, J., Thomas, M., Gissot, L. and Leprince, O.				
TITLE	Differential expression patterns of beta and gamma subunit isoforms suggest the involvement of different SnRK1 complexes in desiccation tolerance, osmotic stress and starvation in germinating seeds of Medicago truncatula				

JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1887)		
AUTHORS	Buitink, J. and Leprince, O.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-MAR-2003) UMR 1191 Molecular Seed Physiology, INRA,		
16 Bd Lavoisier, Angers 49045, France			
Location/Qualifiers			
1. .1887			
/organism="Medicago truncatula"			
/mol_type="mRNA"			
/cultivar="Paraggio"			
/db_xref="taxon:3880"			
/dev_stage="embryo"			
65. .1522			
/codon_start=1			
/product="AKIN betagamma"			
/protein_id="AA061673.1"			
/db_xref="GI:29887975"			
/translation="MYPSLTDVGGSSSSLSSSSSGPIIPKRFVWPVGCTRVYLGISF			
TRMSEHIPMSMEGCPVFQVCSLMPGHQFKFNVDQWRIDEQDFVNGVNGIYNT			
IYIVRRPDILPALISATSSSRSHMEVDNDFGHAEANPRMSDELSRRRIKXFLSE			
HTAYDLPSGKVIALDVALPVKQAEHLVYEQDVSMAPLDFCKSQVGVLSAMDPLK			
ILKELGTGSHLNEQLETHITIAWKEGSKORRALDNNSGPHCFVHAGPKELKD			
VALKLVQNKVSTVPIISLEDGSPQLLRLPNAISGDALSMFYQAKVSSPIVDENSLDLOYSR			
SIPILGTWVNDGPNQGLIRLPNNAISGDALSMFYQAKVSSPIVDENSLDLOYSR			
SDITLAKDKAVARIISLDETNIHQALILGQDANSFYGLNNHGRCHMCLRSLSLHKWE			
RLAKPGVRRLVIVEAGSKRVEGIIISLSDVFRFLIG"			
Alignment Scores:			
Pred. No.: 1 39e-117	Length: 1887		
Score: 1481.50	Matches: 287		
Percent Similarity: 78.57%	Conservative: 65		
Best Local Similarity: 64.06%	Mismatches: 83		
Query Match: 63.26%	Indels: 13		
DB: 8	Gaps: 4		
US-09-357-525C-2 (1-451) x AY247268 (1-1887)			
Qy	3	GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer	22
Db	209	GAGCATATACCATGCTCCGATGGAAGGATGCTCTGTGTTCAGGTATTTGCGAGC	268
Qy	23	LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTrpArgHisAspGlu	42
Db	269	TTGATGCCGGATATCATCATGTTCAATTTAAATGTGGATGTCATGCCGATGATGAG	328
Qy	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu	62
Db	329	CAACAGCAATTTCTAAACGGGAACATATGCGCATAGTAATACTATTTACTTAGTAAGAA	388
Qy	63	TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp	82
Db	389	CCAGATATTTTACCTGCCATCTTAAGTCTGGAACA---TCCAGTAGTCCCATCGAG	445
Qy	83	ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
Db	446	GTGTGATATGATGCTTTGGGCAT-----GCAGAAGCT	478
Qy	103	ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu	122
Db	479	AATCCAAGGATGCGCCATCTGATCTGGAGTCTCTGCGCGCCGCATATCTAAATCTT	538
Qy	123	AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle	142
Db	539	TCGTGAACATACTGCATATGATTTGCTTCCUGAATCTGGGAAGTCATTGCTTGGATGA	598
Qy	143	AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro	162
Db	599	AAATTACAGTTAAGCAAGCATTCATCCATGTTCTTTATGAGCAGGATGATCTATGCA	658
Qy	163	LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu	182

Db 659 CTATGGGATTTTGCAGAGTCAATTTGTTGGAGTCTTAGTGCAATGGACTTCATTC 718
 Qy 183 ILeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis 202
 Db 719 ATATTGAAGAGTTAGGAGTCTCATGATCAATTTGACTGAAGAACCACTTGAGACTCAT 778
 Qy 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTip 222
 Db 779 ACTATAGCAGCCCTGGAAGAGGGAAATCAAGCAACGGAGGCACTTGATAATATGAG 838
 Qy 223 ArgProGlnGlnHis--LeuValHisAlaThrProTyrGluSerLeuArgAspIleAla 242
 Db 839 GGATCAATCTCATTTGTTTTCATGCTGGGCCCAAGGATGCTTAAAGAGCTGGCT 898
 Qy 242 ValLysLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAsp 261
 Db 899 TTGAAGGTTTGTCAAAACAAGGTGTCAACTTCTTCTATCATTT-----TCTTTAGAGGAT 952
 Qy 262 GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
 Db 953 GGTTCATTTCCCAACTGCTGCTGCTTCTTCCATCATCAGGAATTTTAAAGATGATATGC 1012
 Qy 282 ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
 Db 1013 AGGCATTTGAGCACTCTGCTGGTCTTCTTGGCCATTTCTCAACTTCGATTCGATA 1072
 Qy 302 ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321
 Db 1073 CCTTTGGGTACATGGGTGCCTTAAGCTTTGGGGATCCAAATGGACACGCCACTGATAAGGCTA 1132
 Qy 322 ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
 Db 1133 AGCCCAATGCTTCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192
 Qy 342 IleProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr 361
 Db 1193 ATACCAATTTGGGATGAAATGATTCATTTCTTGACATATATTCAGAAGTGATATTACT 1252
 Qy 362 AlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGln 381
 Db 1253 GCATTTGGCTAAGGATAAAGCATATGCTCGGATATCCCTGGATGAAACTTAATATTTCATCAG 1312
 Qy 382 AlaLeuGlnLeuGlyGlnAlaAsnThrProPheGlyPheAsnGlyGlnArgCys 401
 Db 1313 GCATTAATTTGGGACAGATGCAATTTCTCTTATGCTTACAAATGGCAGCAGATG 1372
 Qy 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGlnArgLeuAlaAsnProGly 421
 Db 1373 CATATGTTTGGGTCTGATTCATTCATTCACAAAGTGATGGAGCGCTTGGCTAAACCTGGG 1432
 Qy 422 ValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeu 441
 Db 1433 GTTAGGAGACTTGTGATTTGGAGCTGTGAGCAGCGTGTGAGAGCGGTGGAAGGATATTTCATTA 1492
 Qy 442 SerAspIlePheLysPheLeuLeu 449
 Db 1493 AGTGAATGTTTCAGATTTCTATTG 1516
 RESULT 10
 AC128647
 LOCUS
 DEFINITION Oryza sativa chromosome 3 BAC clone OSUNBb0062G19, complete sequence.
 ACCESSION AC128647
 VERSION AC128647.8 GI:34447243
 KEYWORDS HTG.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 92392)
 AUTHORS Buell,R., Hsiao,J., Zismann,V., Moffat,K.M., Hall,J.,

Gansberger,K., Burgess,S., Jarrahi,B., Shvartsbeyn,M., Brenner,M., Ciocko,A., Pai,G., Vanaken,S., Hansen,C., Utterbach,T., Feldblum,T., Khaliq,H.G., Yuan,Q., Quackenbush,J., White,O., Salzberg,S. and Fraser,C.
 Oryza sativa ssp. japonica cv. Nipponbare OSUNBb0062G19 BAC genomic sequence
 Unpublished
 2 (bases 1 to 92392)
 Buell,R.
 Direct Submission
 Submitted (20-JUL-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 3 (bases 1 to 92392)
 Buell,R.
 Direct Submission
 Submitted (10-APR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 4 (bases 1 to 92392)
 Buell,R.
 Direct Submission
 Submitted (04-SEP-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 On Sep 4, 2003 this sequence version replaced gi:28827864.
 FEATURES
 source
 1. 92392
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /sub_species="japonica"
 /db_xref="taxon:39947"
 /chromosome="3"
 /clone="OSUNBb0062G19"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1e-88 Length: 92392
 Score: 1171.50 Matches: 389
 Percent Similarity: 23.94% Conservative: 34
 Best Local Similarity: 22.01% Mismatches: 25
 Query Match: 50.02% Indels: 1321
 DB: 8 Gaps: 11
 US-09-857-525C-2 (1-451) x AC128647 (1-92392)
 Qy 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 Db 76148 GAACATTTGCCAATGTCTCCATGTTGAAGTGTGCCACCGTTTTCAGGCCATATGCAGT 76267
 Qy 23 LeuSerProGlyIleHis----- 28
 Db 76208 TTGTCCTCCCTGGGATTCATCAGTGGCCCTTTTATCCCAAGTTTATCATGTGTGCTTAG 76267
 Qy 28 ----- 28
 Db 76268 ACTAGCTTTGCTACCAATAAATTATGTTCTGTGATGCAACATCTAATGTGTGAACATT 76327
 Qy 29 -----GluTyrLysPhePheValAspGlyGluTrpArgHisAspG1 42
 Db 76328 CTTTAAATCATGTATATTCAGTACAGTTTGTGTGGATGGGAATGGCGCACGATGA 76387
 Qy 42 uArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgG1 62
 Db 76388 GCGGCAACCTACCAACAGGAGACTGTGTGTGTTAACCTTTATGCTTCACTAGGGA 76447
 Qy 62 uTyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAs 82
 Db 76448 CTTTGACGAGATAAATAATAATTGAGCCCTAGTACACTGGAAGTAGGATGAATGGA 76507
 Qy 82 pValAspAsnGlnAsnPheGlnArgThr----- 91
 Db 76508 TGTGGCAACGACAACTTTTCAACGACTACTTAAGTCCACTTATCTTGTGTTTGAATTTGA 76567
 Qy 91 ----- 91

D	76568	AAATACTTGATAGTAGCTGCTATTATTTTCACAGTTGGGAGAAAACAGTGACACATGTGSTA	76627	D	77647	ATCTTTCAATATCTTTCATGA-ACAGGTGACTATGCTCTCTCTGCTAGTTTCCCTCAA	77705
Q	91	-----	91	Q	155	-----	155
D	76628	TACTTGTGCGCAATTTATCTCTGGTTGATAATAATAGTAGGGTAGTAGTCTTGAA	76687	D	77706	TCCTTCAATCTTCTCTGCTCTGTAGCTGTTCCCTCCAAATCTTCAATAGTTGTTCTG	77765
Q	91	-----	91	Q	155	-----	155
D	76688	GCAATACATGTTGGCAGATACATAATAACACACACTTTAATAGCATTAACAGGTAC	76747	D	77766	ACATTGAACATGATGATATTTTGTCTTGATCTGTAAATATAGTGGAAATAGGTACATACCTT	77825
Q	91	-----	91	Q	155	-----	155
D	76748	ACAAATTTGAAATAACATCTAGCCCATGCCCATCTGTGAGGCAATTTGATGCTACCT	76807	D	77826	TTTAGTAAATATGTTGCCCTGTTGTTGGTACAAATGTTTGGTTTCATTGGCAGCTTCAAT	77885
Q	91	-----	91	Q	155	-----	155
D	76808	GCGAATTTATACATCTTTTCGGTGTACTGAGCAAGTGTGATCAATGAAATGCTGGAGA	76867	D	77886	TCATCTGAGATTGATGAAATGGAACACATGACGCTCATGCTTAAAGTTGCAGCTA	77945
Q	91	-----	91	Q	155	-----	155
D	76868	AGGGGTGAGTTGTAATATACAGCACCTGTGTGCAAAATACAAACGTTCTCTCTCTATT	76927	D	77946	CTTCACCTACTTCACACTTCTTGTACAAACCAATGTTGCATCCTCTCTGTACACATC	78005
Q	91	-----	91	Q	155	-----	155
D	76928	GGAATTTTATTTATAGGAAGCGATACAAAGCTGAGTAGCTGATACATTTATTTTATCT	76987	D	78006	ACCTAGAATCAATGTGTCAACCTTTTAAAGCTTTGAAAGAGCACATATATGTAATAACA	78065
Q	91	-----	91	Q	155	-----	155
D	76988	TCATAAAGTTCAATGTGTGATACATTTATTTTATCTTCACTAATAAAGTTTATGTG	77047	D	78066	TTCACTTAAATTTATTAATCGATACATCAATCAATGGGGTCTAAATTTTGTTAAGATTGA	78125
Q	91	-----	91	Q	155	-----	155
D	77048	TGTGATACATTTCTTTATCTTCACTAABAAGTTCAGGTGTGATACATTTCTTTTA	77107	D	78126	TTAGAGCAAGTAAATTTGTCAGGGTGTGATGATGTCGACACATCCAAATACAACTTA	78185
Q	91	-----	91	Q	155	-----	155
D	77108	CTTCACTAATAAAGTTCAATGTGTGATACATTTATTTTATCTTCACTAATAATTTTCA	77167	D	78186	GAAGAGTTGGGAGTAGGTCTTCTATCTGTTTGCACTCCCAATTTCCATATAGTTA	78245
Q	91	-----	91	Q	155	-----	155
D	77168	GTGTGTATACATTTCTTTTATCTTCACTAATAAAGTTCAATGTGTGATACATTTATTT	77227	D	78246	TAAGAAAGTTCAAGCTCTGTGCTTCTTGTAGCTGCCATCTTGATCTGTGATTTAATTTT	78305
Q	91	-----	91	Q	155	-----	155
D	77228	TTATCTTCACTAATAAAGTTCTGTGTGCGCAATACATTTCTTTTATCTTCACTAACAAGT	77287	D	78306	GGCATCTTGATATAACTGGAATTTTCATAGTTTCTTTAACCATATATTAAGTACAAG	78365
Q	91	-----	91	Q	156	-----	159
D	77288	CATGTGTATTAATAGTCTTCTTCTTGTAGTTTGTGCTTTTGTCTTTTCACTTTTC	77347	D	78366	AAATATTGGCTGTTTGGGAACATACATTAATAAATAGTGTACCTTGGAGGAATCC	78425
Q	92	-----ValThrLe	94	Q	159	oValAlaProLeuTIPAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAs	179
D	77348	TTTCTAGAGCTAAATAATAAATAAACAATTTTCTTTTCTTCTTCAAGCAGGTTCTTT	77407	D	78426	TGTGGCACCTCTATGGATTCATTCAGGGGCCAGTTTGTGGCTTTTGAGCCACATGA	78485
Q	94	sSerAspGlyThrValSer-GluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSe	114	Q	179	pPheIleLeuIleLeuArgGlu	186
D	77408	GTCTGATGCAATTAATTCAGGAAGTCTCTCAGAGAAATTCAGAGCACTATACAGATCTC	77467	D	78486	TTTTTATCTCATATTCAGAGAGGTAATTTATCTGTCACTTACTGTTTGGTGTAGTATT	78545
Q	114	rArgCysArgValSerGluTyrlLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSe	134	Q	186	-----	186
D	77468	TAGGTCTGTGTAGCGGATTTTCTGAATGGACAACTGGGTATGATTTTACTCCAGATTC	77527	D	78546	CTTTAATTTATTTTATATGTGTCATGTTGTTCTCTATTGTATTTTATTATTGTTACAA	78605
Q	134	rGlyLysVal	137	Q	186	-----	186
D	77528	TGCAAGGT-ATCGAAATAGCTCGCTCTGTATGAATTTTGTGCTCTAAATGTTATTC	77586	D	78606	GTGATGGTTTAAATGTTCAATTCAAATATGCTGCCACTTTTAGTGGCCCTCGAAACGAAA	78665
Q	138	-----IleAlaLeuAspIleAsnLeuProValLysG	148	Q	186	-----	186
D	77587	TTATTACATTAATTTACAAATTCAGGTCAATGTCTAGACGTTAATTTGCTGTGAAGCA	77646	D	78666	TTGCGCAATCTGTTTCTTTTAAAAAATCTGAAAGCAGTGTAAATATCATCTTTG	78725
Q	148	nSerPheHisIleLeuHisGlu	155	Q	186	-----	186
				D	78726	ACAGTATTAAGTTATTGACCTAGAGCTTATTTCATCAATGAATTAATAAGACGACTGTTTC	78785

Qy	237	LeuArgAspIleAlaValLysLeuGlnAsnGlyIleSerThrValProValIleTyr	256	Qy	79866	CTGAGGAAATTCGTATGAAGATCTACTGCAAAACGGTGTCTTCTACCGTTCCTCAATATATGTTT	79922
Db				Db			
Qy	257	SerSerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIle	276	Qy	79926	TCCTCATCACAGATGCTCATATCCCAATTTGTTGCATCTTGCTTCCCTTCTCGGAATT	79985
Db				Db			
Qy	277	LeuLysCys	279	Qy	79986	TTGAAATGTAAGTGTCTATGTTATGTCATGAAAGTGTGTCTATACACAACATATTAC	80045
Db				Db			
Qy	280			Qy	80046	ATGATAGCTTAACAACAATTTCAITTTCCAGGTATTTGTAGATATTTTAAAAAATTTCTCAAGG	80105
Db				Db			
Qy	289	YasnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerIrpValProLy	309	Qy	80106	TAATTTACCTATTTTGAGCCAACTGATGACAAATTCCTCTGGGTACCTGGGTCCAAA	80165
Db				Db			
Qy	309	sileGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerSe	329	Qy	80166	ATTGGTGTATCTAATGCGCTCCATTGGCTATGTTGGCGCTTAACACATCTCTTAGCGC	80225
Db				Db			
Qy	329	rAlaLeuAsnMetLeuValGln	336	Qy	80226	TGCCCTGAATTTGCTGGTTCAAGGTATGTTATCTGCTGTTTTTACCATAACCAANGCTG	80285
Db				Db			
Qy	336			Qy	80286	GGAAAGTACTTGGCAATTTCTGTTCTTGGTGCAATTACTGAGTTTAATACTAATCCACT	80345
Db				Db			
Qy	337			Qy	80346	GAGATCTCTTACAGCTTTCGATTATTTACGCTATCTTGTTCGCCAAATTAATGCTTATTA	80405
Db				Db			
Qy	346	PAspAsnAspSerLeuLeuAspThrTyrSerArg	357	Qy	80406	CAATTAATCTCTACTCTAAATTTGTTTCATTCAGCTGGTGTGAGCTCAATACCAATTGTGA	80465
Db				Db			
Qy	358	erAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetT	378	Qy	80526	TTTGTGTAGTCATATACTATTTCTTTGTTCATGTTATTCATTGTGTTTTTCTCCATTATGTA	80585
Db				Db			
Qy	378	hrIleHis	380	Qy	80586	GTGACATCACAGCTCTTGCAAAAGACACAGGTGTACACACATTCGCGCTAGTAGAGTGA	80645
Db				Db			
Qy	380			Qy	80646	CCATTCATCAGGTATGCTAGTAATAATTTAATGCGTCTTGATTGTTGTGCTGGAAT	80705
Db				Db			
Qy	380			Qy	80706	TTTACTCGGATACAGTTGTACAATAATTTTTCGAGAGTTCCTTTTGTGCTTTGAAGACTG	80765
Db				Db			
Qy	380			Qy	80766	GTTAACTTAGAAAAATAAATCTTTACTGATTTATGATATGATAGAGGAAATCTGTATCGA	80825
Db				Db			
Qy	380			Qy	80826	TACTGCAATCACACATATATACTATGAGGACATGCTTTGTGGTCTTGCCATTATTATGAC	80885
Db				Db			
Qy	380			Qy	80886	CCTGTGCTTTTTCGACGACGACATGGAGCATGCACACATCTAAAAACAAAGCTGTA	80945
Db				Db			
Qy	380			Qy			


```
/codon_start=1
/product="putative lipase/acylhydrolase"
/protein_id="AAP21383.1"
/db_xref="GI:30102970"
/translation="MTMKKQKQNMKLGILLVAVAVAMAPAGPASFSSGRGAPMPAV
IVGDSIVDPGNNNLKTQIKAHAPYGMDFANSEPTGRVSNGLIPTDFIVQGLNVKQ
LMPPTLGVLSPELDKGTGSPAGATGYDPLTVIVSVLTLDQGIYEVEHYRKLGVG
VGEETARIGDALFVWCAGTDIDANTFYFTTPRSVEYDIPSYVDLLVSSAAKLDOV
AALGARIGFVGLPPTIGCVPSQRTILGGPHRRCEERENYAAKLFNSRMBEVIKKNP
ATRMVVDIYITLQELVENDKYGFTETTRGCCGTTEBVTGLCDARFVDICDNVSN
HVFDSVHTQRAYKIIVDYIFDNVLQFLIA"
repeat_region
6410..6437
/rpt_family="AT-rich"
781..7804
/rpt_family="AT-rich"
/rpt_family="AT-rich"
12122..13806
/rpt_family="AT-rich"
/notes="OSJNBa0059G06.3"
join(e12122..12362,12851..13020,13702..13806)
/notes="predicted by fgenesh"
/notes="OSJNBa0059G06.3"
/notes="OSJNBa0059G06.3"
/notes="OSJNBa0059G06.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP21385.1"
/db_xref="GI:30102972"
/translation="MRVRFKVGASPGFKESVSGACWNGAPRAGDERPPGADGNG
SEAMPIDGGARKWFGVIGSSRTVFARVVDLGGDDARKNMAMWMAKAAVDQTAEP
SSRECLARGSSSESQPCSELAAMRQSMVTVAEVELDNGNREESRDGFKVDSIRGIG
TSIWRPKGSD"
repeat_region
13321..13364
/rpt_family="CA-rich"
15008..15049
/rpt_family="TCn"
16325..16352
/rpt_family="TTCCn"
16964..17005
/notes="Overlap BAC OSJNBa0033P04 has 49 TA repeats in this
region."
repeat_region
16964..17005
/rpt_family="TA"
17177..17203
/rpt_family="AT-rich"
17384..18823
/notes="OSJNBa0059G06.4"
/notes="contains Pfam profile PF01554 (MatE)"
<17384..>18823
/notes="OSJNBa0059G06.4"
17384..18823
/codon_start=1
/product="putative MatE efflux family protein"
/protein_id="AAP21388.1"
/db_xref="GI:30102975"
/translation="WTPPPPPPPHKKTKWAESVASEFRAQRIAPLIAMNLTWFAKL
AVTAFLEGLDQLAGTGFANVTGFVLTGLCAAMDPICGOAGHANGKLLRK
TLVMTATLIGASIPATLMLHVDALLRFQOADMSSNARSVYVCLLPDLAVTSFN
PLKXVLSQRQTPLTPASALALAHVPLTWMTATRGIGVATVAVVSDLAVALVLA
GVLVSRKKRAGGGGWEQTRGEVRLRLRAVPSCLNCLCWCYVTLVLTGRLP
DARTVAVMTVINFEDYLLFAGMLSLVSASVRSNELGAGHMAARAGMVSIVGGA
VGVGGVAVMAVARAWGSIYSSDAGVREGVGRVMEVAVLEVVNPFNVCGGIVRG
ARPVGMVAVVAGFVYVLLPLGVLAFAKRLGIQLLGLFLVGAASLAVLTFIARM
DKPAQAQARTTATVAQFHQHVWQ"
repeat_region
18325..18515
/rpt_family="G-rich"
20002..21960
/notes="OSJNBa0059G06.5"
join(<20002..20123,20242..20298,20375..20530,20579..20654,
20811..20903,21132..21245,21366..21960)
/notes="OSJNBa0059G06.5"
join(20002..20123,20242..20298,20375..20500,20579..20654,
20811..20903,21132..21245,21366..21419)
```

```
/gene="OSJNBa0059G06.5"
/codon_start=1
/product="unknown protein"
/protein_id="AAP21391.1"
/db_xref="GI:30102978"
/translation="MAPPPCCVUVRCCSGSLSPPPPPPPPPSPPSNNLRIARREFV
NRSSELAATLAIIFHLSGTXPYLGVSQPSLALCPATNMCVSTCEIDTISIHVAPW
LNPQKORAKPITKHEAINQLIQVVTQTPDNFTPLRVEKTDYDVRVEYSPFIFGV
DDVEFFPPGNKSIQVRSASRSGFIDFNANKRVKALRLALENKGWASBSTI"
repeat_region
20063..20084
/rpt_family="(CCG)n"
20673..20703
/rpt_family="(TTTTC)n"
22163..22193
/rpt_family="AT-rich"
24255..25676
/notes="OSJNBa0059G06.6"
/notes="predicted by fgenesh"
Alignment Scores:
Pred. No.: 1.81e-88 Length: 145115
Score: 1171.50 Matches: 389
Percent Similarity: 23.94% Conservative: 34
Best Local Similarity: 22.01% Mismatches: 25
Query Match: 50.02% Indels: 1321
DB: 8 Gaps: 11
US-09-857-525C-2 (1-451) x AC096690 (1-145115)
QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 138916 GAACATTGGCAATGTCTCTGTTGAAGGTGTCCACCGTTTTCAGGCCATATGCAGT 138957
QY 23 LeuSerProGlyIleHis----- 28
Db 138856 TTGTCCTCCCTGGGATCACTAGTGCCTTTTATCCCAAGTTTATCATAGTGTCTTAG 138797
QY 28----- 28
Db 138796 ACTAGCTTTGCTACCCATAAATTATGTTTGTGTGATGCAACATCTAATGTGTGAACATT 138737
QY 29-----GluTyrLysPhePheValAspGlyGluTyrArgHisAspG1 42
Db 138736 CTTTAAATCATGTATATTCAGTACAGTTTGTGTGGATGGGATGCGCACGATGA 138677
QY 42 uArginProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgG1 62
Db 138676 GCGGCACTACCATACAGGAGACTATGTTGTGGTTAACACTTTTATGCTTACTAGGGA 138617
QY 62 uTyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAs 82
Db 138616 CTTTGACCAGATAAATAACAATATTGAGCCCTAGTACACCTTGAAGTAGGATGACATGGA 138557
QY 82 pValAspAsnGluAsnPheGlnArgThr----- 91
Db 138556 TGTGACCAACGACAAATTTTCAACGTAAGTGCCACTTATCTGTTTGTAAATTTGCA 138497
QY 91----- 91
Db 138496 AAATACTTGATAGGATAGCTGCTATTTTTCACAGTTGCGAGCAACAGTGACACTGTGGTA 138437
QY 91----- 91
Db 138436 TACTTGTGCGCAATTTATCTCTGGGTGATAATATGTAGGGGTAGTATGCTTCTGAA 138377
QY 91----- 91
Db 138376 GCATACCATGTTGGCGCAGATACATATACACACTTTTAAATAGCATTACAAGGTCAAC 138317
QY 91----- 91
Db 138316 ACAATTTGAAATAAACAATCTAGCCCCCATCTGTGACGCCATCTGTGACGCATTTGTATGCTACCT 138257
```

```
QY 91 ----- 91
Db 138256 GCGAATTTATACATCTTTTCGGTGTACTAGCAAGTGTGATCATTTGAATGCTGGGAGA 138197
QY 91 ----- 91
Db 138196 AGGGTGGAGTGTAAATTATCAGCACCTGTGTGCAAAATACAAATGTTCTCTCCTCTATT 138137
QY 91 ----- 91
Db 138136 GGAATAATTATTTATAGAAAGCATACAAGCTGAGTAGCTGTACATTTATTTTATCT 138077
QY 91 ----- 91
Db 138076 TCACATAAAAGTTTCATGTGTGTGATACATTTATTTTATCTTCTACATAAAGTTCATGTG 138017
QY 91 ----- 91
Db 138016 TGTGATACATTTCTTTTATCTTCACTAAAGATTCACGTGTGTGATACATTTCTTTTA 137957
QY 91 ----- 91
Db 137956 TCTTCAATAAAGTTCATGTGTGTGATACATTTATTTTATCTTCACTACAATGTTCAAT 137897
QY 91 ----- 91
Db 137896 GTGTGTGACATTTCTTTTATCTTCACTAAAGATTCACGTGTGTGATACATTTTATT 137837
QY 91 ----- 91
Db 137836 TTATCTTCACTAAAGTTCGTGTGTGCAATACATTTCTTTTATCTTCACTACAAGGT 137777
QY 91 ----- 91
Db 137776 CATGTGTGTTTAAATGCTTCTTATTTGTTTAGTTTGGCTTTTGCCTTTTACCTTTC 137717
QY 92 ----- ValThrLe 94
Db 137716 TTCTTAGAGCTAAATATATAATAACAACATTTCTTGTCTTCTCAAGCAGGTTTCTTT 137657
QY 94 nSerAspGlyThrValSerGluGlyThrLeuArgValSerGluAlaLeuGlnIleSe 114
Db 137656 GTCGTAGGCAATTATTCAGGAAGTCTCTCAGAAATTCAGAGCAGCATACAGATCTC 137597
QY 114 rArgCysA-gValSerGluPyrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSe 134
Db 137596 TAGTGTCTGTGTAGCGAATTTCTGAATGCACAACTGGGTATGATTTATCCAGATTC 137537
QY 134 rGlyLysVal ----- 137
Db 137536 TGSCAAGGT-ATGGAAAATGAGCTCGTCTTGTATGAATTTTAGTGCTCTTAATGGTTATTC 137478
QY 138 ----- IleAlaLeuAspIleAsnLeuProValLysG1 148
Db 137477 TTATTACATATATTTACAAATTCAGGTCAITGCTCAGAAATTCAGAGCAGTAAATTTGCCGTGAAGCA 137418
QY 148 nSerPheHisIleLeuHisGlu ----- 155
Db 137417 AUCTTTTCAATTTCTTCATGA-ACAGGTGACTATGCCCTTCTCTGCTAGTTTCCCTCCAA 137359
QY 155 ----- 155
Db 137358 TCTTTCATTTCTTCTCTGCTCTGTAGCCTGTTTCCCTCCAAATCTTTCATTTAGTTGTTCTG 137299
QY 155 ----- 155
Db 137298 ACATTGAACACTAGTGATATTTTGTGCTGATCTGTAATAATATAGTGGACAGGTACATACCTT 137239
QY 155 ----- 155
Db 137238 TTTAGTAATAATATGTTGGCTGTTGTTGGTACAAATGTTTGGTTTCATTTGCCAGCTTCAAT 137179
QY 155 ----- 155
```

```
Db 137178 TCATCTGAGATTGCATGAAATGGAACACACTGCAGCCTCATGCCCTTAAAGTTGCAGCTA 137119
QY 155 ----- 155
Db 137118 CTTCACTTACTTTCACACTTCTTGTATACACCAATGTTGCATCCTTCTCTTGTATACACATC 137059
QY 155 ----- 155
Db 137058 ACCTAGAAATGCATGTGGTCAACCTTTTAAAGCTTGAAAGAGCACATATATATGTATAAAAAACA 136999
QY 155 ----- 155
Db 136998 TTCACTTTAAATTTATTAATCGATACATCATCAATGGGGTCTAAATTTTGTTHAAGATTGTA 136939
QY 155 ----- 155
Db 136938 TTAGAGCAAGTAATTTGTCAAGGTGATGATTGATGTGCAGACATCCAAATACAACAAATTA 136879
QY 155 ----- 155
Db 136878 GAAGAGGTGGAGGAAGTAGGTTCTATTTCTGTTGCACTCCACAAATTTCCATATAGGTTA 136819
QY 155 ----- 155
Db 136818 TAAGAAAGTTTCAGCTCTGTGCTTCTTAGCTGCCATCTTGATTTCTGATTTAAATTAATTTT 136759
QY 155 ----- 155
Db 136758 GGCATCTTGATATAACTGGAATTTTCATAGTTTCTTTAAACCATATATATTGAAGTACAAAG 136699
QY 156 ----- GlnGlyIlePr 159
Db 136698 AAATATTGGCTGCTTTTGGGNAACATACATTAATAATATAGTGTACCTTGCAGGGAAATTC 136639
QY 159 oValAlaProLeuPheAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAs 179
Db 136638 TGTGGCACCTCTATGGGATTCATTCAGGGGCCAGTTTGTGGCCCTTTTGGACCCCACTGGA 136579
QY 179 pPheIleLeuIleLeuArgGlu ----- 186
Db 136578 TTTATATCTATATTTAGAGAGGTAAATTTATCTGTCACTTACTTGTGTTGGTGTAGTATT 136519
QY 186 ----- 186
Db 136518 CTTTAAATTTTATTTATGTCATTTGTTGTTCTCTATTTGTTATTTTATTTATTTTATTTTACAA 136459
QY 186 ----- 186
Db 136458 GTGATGGTTTAAATGTTTCATTTCAAATATGCTGCCACTTTTAGTGCCCTGCAGAACGAAA 136399
QY 186 ----- 186
Db 136398 TTGCGCAATTTCTGTTTTTTTTTTTTTAAATAAATCTGAAGCAGTGAATATCATTTCTTTG 136339
QY 186 ----- 186
Db 136338 ACAGTATTAAATTATTTGACCTAGAACCTTATTTACATGAATATATAAAGACGACTGGTTC 136279
QY 186 ----- 186
Db 136278 ATTAATGGAATGCGGAGTCTATGCATTTGAACAACTGACTGCTTGTAAATCCGATTTCT 136219
QY 186 ----- 186
Db 136218 GCTGTATCACTTAGTAACCATGCATCTCATGGAATTTGTTGGGTGACCACAGAAGTT 136159
QY 186 ----- 186
Db 136158 CAGCCAAGATAGTATCATGTTGAACCTGTTTTCCCTTAATTTGGAAGTGTTCAGTTAGTA 136099
QY 186 ----- 186
```

Db 136098 GAACCTCCAAAATTTGAAATCGTATCTTTTGGAGGTGGACTCTAGAAAATAGAAATAA 136039
 QY 186 ----- 186
 Db 136038 ATATGCTAGTAGAGCTGCATCATGATACCTTTATTTATGCTGCTTTGAATGCTTC 135979
 QY 186 ----- 186
 Db 135978 TGGTTCAGGTAAAGGATTAATACTCGACCAATGATTTGTACTGGAACAATTTGT 135919
 QY 187 ----- LeuGluThrHisG 191
 Db 135918 GAGCATATTCACACAGTAGCATTAATATCTTGTTATTTGACGCTGGAATCATG 135859
 QY 191 lySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSerAlaTrpLysGluAlaL 211
 Db 135858 SCTCCAACTCTGACAGAAGAGCAGCTTGAACACATACATATCTGGTGGAGAGGCCA 135799
 QY 211 yArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLeuValHis- 230
 Db 135798 AGCGTCANAATTATGCCAGAAACGAGGGTTCTTGGAGGGCAATACCACTTATGTTTCATG 135739
 QY 230 ----- 230
 Db 135738 TGAAGTGTCCACTGTGTTTGGAAACCAATTGCAATCTGTATCGCACTCACCACAAATTGA 135679
 QY 230 ----- 230
 Db 135678 GACTGCTGCCATCAACTTCTATCTATATTTACTGGCCAAATTTACTCCAGTGAATTTA 135619
 QY 230 ----- 230
 Db 135618 GAATTCAGTATTCATTCATGAGTAAATATTTTACCTTTGATTTATTTGCAATTTGACA 135559
 QY 230 ----- 230
 Db 135558 AATTGTTGTCAACAATTTGTTGTGTGAGTACCACATAACAGTTTITGGCACACCAATGACAC 135499
 QY 230 ----- 230
 Db 135498 ACATTTTGTCTAACATTTTTTGGCACATGATGACACATTTTACTTTTGTGTACAAAC 135439
 QY 230 ----- 230
 Db 135438 TAAGTTACTACTGGATGCTGAATATATTTTGTCTTAATTTAGCTTTAATTTGCCATGCTCA 135379
 QY 230 ----- 230
 Db 135378 AGTGAGTTACTTAGTGAACTGCAATTTGTTTGTGTGAAATTTTCATCGAACCCCTTGCAAT 135319
 QY 230 ----- 230
 Db 135318 CAAGTGTCTCTTTTCATTTCCATCCTAGTATTTTATTAACATGCTTAGTTTATATGATC 135259
 QY 231 ----- AlaThrProTyrGluSer 236
 Db 135258 ATGTGCTTCATGCACTATGCTGTGAGTGAATCTATTTGTTAGGCCACCCCTTATGAATCC 135199
 QY 237 LeuArgAspIleAlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyr 256
 Db 135198 CTGAGGGAAATTTGCTATGAGATACCTGCAAAACGGTGTCTTACCGTTTCCAAATATATGTT 135139
 QY 257 SerSerSerAspGlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIle 276
 Db 135138 TCCTCATCACCAGATGGCTCATATCCAAATTTGTGCAATCTGTCTTCCCTTCTGGAAAT 135079
 QY 277 LeuLysCys ----- 279
 Db 135078 TTGAAATGTAAGTGTCTTATGTTATTTGCAAGAAAGTGTGCTATACACAAATATTTAC 135019
 QY 280 ----- IleCysArgTyrPheLysAsnSerThrG1 289
 Db 135018 ATGATAGCTTTAAACAATTTTCATTTTTCAGGTAATTTGTAGATATTTTAAATAATCTCTCAAG 134959

QY 289 yAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpValProLy 309
 Db 134958 TAATTTACCTATTTTGGCCACCTGTATGCAATTCCTCTGGTACCTGGGTCCAAA 134899
 QY 309 sileGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerSe 329
 Db 134898 AATTGGTGATCTTAATGGCGCTGCATTTGGCTATGTTGGGCGCTAAACACATCTCTTAGCGC 134839
 QY 329 rAlaLeuAsnMetLeuValGln ----- 336
 Db 134838 TGCCCTGAATTTGCTGGTTCAGAGTATTTATCTGTCTGTGTTTACCATAACCAATGCTGT 134779
 QY 336 ----- 336
 Db 134778 GGAAGTACTTGCATTTCTGTTTCTTGGTGCATTTACTGAAGTTTAACTAATTCCTACT 134719
 QY 336 ----- 336
 Db 134718 GAGATCTCTTACAGCTTTTCGATTATTTACGCTATCTTGTGTCCTTGTGTCCTTTATTA 134659
 QY 337 ----- AlaGlyValSerSerIleProIleValAs 346
 Db 134658 CAATTACTCTGTACTCTAATTTGTTTCAATCAGCTGGTGTGAGCTCAATACCAATTTGGA 134599
 QY 346 pAspAsnAspSerLeuLeuAspThrTrpSerArg ----- 357
 Db 134598 TGACATGATCTCACTGCTTGCACATATTCAGAGAGGTGCCATTTTGATACTACCATTC 134539
 QY 358 ----- S 358
 Db 134538 TTTGTGTAGTCATATACTATTTTGTGTCATGTAATTCATTTGTTTCCATATGTA 134479
 QY 358 exAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetT 378
 Db 134478 GTGACATCAACAGCTCTTGCAAAAGACAGGTGTACACACATTCGCTAGATGATGA 134419
 QY 378 hrIleHis ----- 380
 Db 134418 CCATTCATCAGGTATGCTAGTAAATATTTAATGCTCTTGTATTTAGTTTGTCTGSAAT 134359
 QY 380 ----- 380
 Db 134358 TTTAGCTGGGATACAGTTGTACATAATTTTTCAGAGATTTCTTTTGGTCTTGAGACTG 134299
 QY 380 ----- 380
 Db 134298 GTTAACTTAGAAAATAATACTTTTACTGATTTATGATATGATAGGGAATCTGTATCGA 134239
 QY 380 ----- 380
 Db 134238 TACCTGCATCACATATATATATGAGGACATGCTTTGTGTTCTGGCATTTATTAGC 134179
 QY 380 ----- 380
 Db 134178 CCTGTGCTTTTTCACGACAGAGCATGAGCATGSCACACATCTAAACAAAGGCTGTA 134119
 QY 380 ----- 380
 Db 134118 GGCTGTAGTTGATGTTGTTTGTCCATTTATCACTTCTTACCTTTCCATACCGGTCAAACA 134059
 QY 380 ----- 380
 Db 134058 ACATATATAATAAGATAAGTAGTGAATGCTGTTTAAAGACGGCTTTCTCTGTGAAT 133999
 QY 380 ----- 380
 Db 133998 TGTGTATATGTTATCCATTCGAGCTGGTGTGATTTTGTGTGAACCTACTCATTTTGT 133939
 QY 381 ----- GlnAlaLeuGlnLeuGlyClnAspAlaAsnThrProPheGlyPheAsnG 398
 Db 133938 GTTATGTTCAAGCTCTGACGCTTGGACAGGATGCAATTTCCGCCCTTTTGGATTTCTTAACG 133879

QY 398 lyGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuA 418
 Db 133878 GGCAGAGATGTGAGATGTCTCCGGTCTGACACITTTGTGTGAAGTGTGGAGCGATTAG 133819
 QY 418 laAsnPro----- 420
 Db 133818 CTAATCCGTGTAATTAATCTGTGCTAGATATGCTTCTCGAAGCAACTTTTCGGGTCAA 133759
 QY 421 -----GlyValArgValPheIle 427
 Db 133758 CTACTAATAATTTTTTTTTCATTTTATTTGTTGTTAGGGTGTGGCGCTGTTTTCATT 133699
 QY 428 valGluAlaGlySerLysArgValGluGlyIleLeuSerLeuSerAspIlePheLysPhe 447
 Db 133698 GTGGAAGCTGGTAGCAACGCTGGAGGGGATAATATCACTGAGTGATATTTTCAAGTTC 133639
 QY 448 LeuLeuSer 450
 Db 133638 TTGCTGAGC 133630
 RESULT 12
 LOCUS BT009623 982 bp mRNA linear PLN 20-JUN-2003
 DEFINITION Triticum aestivum clone w.reln.pk0143.e2.fis, full insert mRNA
 sequence.
 ACCESSION BT009623
 VERSION BT009623.1 GI:32129174
 KEYWORDS FLI CDNA.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Triticum.
 1 (bases 1 to 982)
 Tingey,S.V., Walters,P., Powell,W., Dolan,M., Miao,G.-H.,
 Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
 Direct Submission
 Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and
 Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA
 FEATURES
 source Location/Qualifiers
 1..982
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="w.reln.pk0143.e2.fis"
 ORIGIN
 Alignment Scores: Length: 982
 Pred. No.: 2,02e-82 Matches: 205
 Score: 1068.00
 Percent Similarity: 95.61% Conservative: 13
 Best Local Similarity: 89.91% Mismatches: 10
 Query Match: 45.60% Indels: 0
 Db: 8 Gaps: 0
 US-09-857-525C-2 (1-451) x BT009623 (1-982)
 QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrShuSerLeuArgAspIleAlaVal 242
 Db 11 AGATCAATCAGCACTAGTATGATGCACCCCTATGAACTCTTGAGGGGTATTGCCATG 70
 QY 243 LysLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGly 262
 Db 71 AAAATACCGAACTGGCATTTCTACAGTCCCAATCATCTATTTCATCGTCATCAGATGA 130
 QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
 Db 131 TCGTTTCCGAGCTGTGTGATCTGTGATCCCTTTCAGGAATTTTGAATGATCTGTAGA 190
 QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 Db 191 TACTTCAAGAACTCCACTGGTAGTTTGGCGATTCTTAACCAACCAAGATGCTCAATTCOG 250

QY 303 LeuGlySerTyrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
 Db 251 CTGGTACCTGGGTTCCAAAATTGGTGAACCAATGGTCAATCCATTGGCTATTGCGG 310
 QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
 Db 311 CCTAATACATCTCTTAGCTCTGCCCTTAACCTTGTGGTCAAGCTGGAGTTAGTTCATA 370
 QY 343 ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 Db 371 CCCATTGTGATGATACAGACTCGCTGATGACACACTATCCAGAGTGATCATCACAGCT 430
 QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
 Db 431 CTAGCGAAGACAGTCTACACCCATATCCGCTAGATGAGATGACCATTCATCAGGCC 490
 QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGln 402
 Db 491 TTGCAGCTCGGGCAAGACGCGAATTCACCTTTTGGACTTTTCAATGGTCAAGAGATGCCAG 550
 QY 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
 Db 551 ATGTGTCTCCAGCTGACCCCTTTGCTGAAGTTATGGAGATTTGGCTAATCTCGGGTG 610
 QY 423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleLeuSerLeuSer 442
 Db 611 CGTGGGTGTTTCATCTGAGGCTGGCAGCAAGCGAGTGAAGGCGTAAATATTCGCTGAGC 670
 QY 443 AspIlePheLysPheLeuLeuSer 450
 Db 671 GACATATTAAGTTGCTGCTGAGC 694
 RESULT 13
 LOCUS F7G19 110514 bp DNA linear PLN 16-SEP-1997
 DEFINITION Sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1,
 complete sequence.
 ACCESSION AC000106
 VERSION AC000106.1 GI:2342673
 KEYWORDS HTG.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 110514)
 Osborne,B.I., Vysotskaia,V.S., Toriumi,M., Yu,G., Oji.,O.,
 Buehler,E., Conway,A.B., Conway,A.R., Dewar,K., Feng,J., Kim,C.,
 Kurtz,D., Li,Y., Shinn,P., Sun,H., Davis,R.W., Ecker,J.R.,
 Federspiel,N.A. and Theologis,A.
 The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1
 Unpublished (1997)
 2 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (20-JAN-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 3 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (17-FEB-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 4 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (27-FEB-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 5 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (03-MAR-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA

gene	HKAGAGMGLGILGAGLLGGLLNGEAYSDIADMGDMGDMGDFP*	Db	39908	GTGGAATATGAGATGGAT---GATGTCTTCTGCGAAGGTAAGAA--TCTGACCTTTC	39963
CDS	17933..21752	QY	98	rValSer-----	100
	/gene="F7G19.7"	Db	39964	CTTTTCCAAATGACAGTGTATTTCTCGCTGCTTGATGTTCCAGGTTTCCGAGTTGCTT	40023
	/join(17933..18538,18639..18781,18879..19029,19082..19181,	QY	100	-----	100
	19238..19367,19470..20115,20194..20523,20583..20618,	Db	40024	CCAGTTTGGTTCTTTAAATTTTGGATCCGTGGTGGTATCCTTCAGATTATTTTAACTGT	40083
	20661..21032,21125..21194,21284..21369,21470..21577,	QY	100	-----	100
	21678..21752)	Db	40084	TTGTCAATAGTTCAAGTATAATTTAGACCAATTTATCTTAATGTATAAGCAGGA	40143
	/gene="F7G19.7"	QY	100	-----	100
	/note="Similar to Vicia sativa ENBP1 (gb X95995)."	Db	40144	TTCCTTAAGTTCCATTAATTTGTCCTGAGACTACATTCAAATTTAGTTTCTGACTCTT	40203
	/codon_start=1	QY	101	-----	108
	/evidence=not experimental	Db	40204	TCACATGCTGATTATATTGTAGGCTGACCGCTCCAGGAAGCTGTTCTTAGGATGTCGG	40263
	/protein_id="BAB70402.1"	QY	108	uAlaAlalleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTy	128
	/db_xref="GI:2342679"	Db	40264	GSTTGATTTGGAGCTTCTCGTCACCGCATATCGTTTATTTGCAACCCGCACTGCATA	40323
	/translation="MQVNEDETCDVIRNANEQTSANGIGNGESIPGIPDDLRC	QY	128	rAspLeuLeuProAspSerGlyLys-----	136
	KESDGQRKRTAMSWADKTVCEKHQIQAKRAANSAPFRANQKAKRSLSIGTDYISE	Db	40324	TGAGCTGCTCCAGATCGGCAAGTTGGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	40383
	GRQDDFELPVTSIDHYNGLASAKSNGLEKRNKSLMRYSPETPMRSFSPRAVD	QY	136	-----	136
		Db	40384	AATTTCTGTCTTCATGCTCTATAGATTATCATAGTAAGTAAGTAACCTTTGGTCTTCTACTCG	40443
		QY	137	-----	146
		Db	40444	GAAATGCTCAAGGTGANTGGAAATTTTACAGGTTATTCGATTCGATTCGATTCGATTCG	40503
		QY	146	allysGlnSerPheHisIleLeuHisGlu-----	155
		Db	40504	TGAAGCAGATTCCTCATATATCTATGAGCAGGTTTACTTATTTGCAATTACCTTGT	40563
		QY	155	-----	155
		Db	40564	GTGAGCTTCTGTTCTGTTAGGATGTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	40623
		QY	155	-----	155
		Db	40624	CTAATATTGTTGCATCCAGACCCGGAAGGAAGGAAGTAAGTAAGTAAGTAAGTAAGTA	40683
		QY	156	-----	156
		Db	40684	CGTCTAATATATCTTCGAGATTGTTGGCTTTCTGCTTTCTGCTAATTCAGATGCTTTC	40743
		QY	156	lnGlyIleProValAlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuS	176
		Db	40744	AGGAAATCCCTTTGGCTCTCTTTGGGACTTTGGAAAGGCGCAATTTGTTGGAGTCTTG	40803
		QY	176	erProLeuAspPheIleLeuLeuLeuArgGlu-----	186
		Db	40804	GTCCACTAGACTTCATTTCTAATCTGAGAGAGGTGGCTTCCTNTTAAACTGTTCTTTC	40863
		QY	186	-----	186
		Db	40864	TAGACATTCCATCGAGTTACTATTTGATATTTCTTAGAATTTCCAGAGGAGTCTGGA	40923
		QY	186	-----	186
		Db	40924	TATCTCATAAATCACATAGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATG	40983
		QY	187	-----	196
		Db	40984	TTGAGTATATAGCAAAATTTGTAATTTGAGCTTTGGAACTCATCGATCAACTTGACAG	41043

Alignment Scores:	3.48e-74	Length:	110514
Pred. No.:	1003.50	Matches:	306
Score:	29.30%	Conservative:	65
Percent Similarity:	24.17%	Mismatches:	75
Best Local Similarity:	42.85%	Indels:	823
Query Match:	8	Gaps:	15
DB:			
US-09-857-525C-2 (1-451) x F7G19 (1-110514)			
QY	3	GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer	22
Db	39254	GAGCATGCTCCATATGTCACCGCTCGAGGGTGCCTACTGTTTTCATGATCAATTCGAC	39313
QY	23	LeuSerProGlyLeu-----	27
Db	39314	TTGACGCCAGGATATCATCAGGTTAATAATCTTGTACTCGACCTGATCATGCGCAATTC	39373
QY	27	-----	27
Db	39374	TTTGTTAATTAATCCCTGCGAGAGCATTTATGCTTGTAGTATTTAGCGATGCCATA	39433
QY	27	-----	27
Db	39434	ATTGCTCTTCAATTTAGTCTCTGATACTTAACCTATGCAATGAGATCTTATATACCTCT	39493
QY	27	-----	27
Db	39494	GTCAATATTGATGACTGTGGCTGAAATTTGCAATTTCTGCTAATTCATCAAAATTT	39553
QY	27	-----	27
Db	39554	GACTTATATACATCCATCCACATACAAAGTTGGGCTCTTCATGCTGTGATACCTA	39613
QY	27	-----	27
Db	39614	AATCACTTCTTTTAATAAGGCTACAGTGTGTATAGATTCATATATTTGGCAGTT	39673
QY	27	-----	27
Db	39674	GGATGACTATATATACAAAGATTTAATTAATTAACAAATGAGATGCTTAAGAAGCCACT	39733
QY	28	-----	38
Db	39734	AATGTTACATTAATGTTATTTCAATCTCATGCAATGAGTTTCTTGTGATGGGAATG	39793
QY	38	pArgHisAspGluArgGlnProThrIleSerGlyGluPheGlyIleValAlaThrLeuTy	58
Db	39794	CGGCGACGATGAGCACCAACCATTTTGAAGCGGAAATGGTGAGTAGTGAATACATATTT	39853
QY	58	rLeuThrArgGluThrAsnGlnIleAsnThrLeuSerProSerThrProGlySerAr	78
Db	39854	TATAACTGGGCCA---GATATGGTTCTCGCTGGTTTATGACCGACGACACTGGGC---CG	39907
QY	78	gMetAsnMetAspValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyTh	98

QY 196 luGluGlnLeuGluThrHisThrIleSerAlaTrpIysGluAlaLysArgGlnThrAsnG 216
|||||
Db 41044 AGGAGAGCTTTGAGCGCACACATAGACCTCTGGAAAGAGGGAAGGCTCATATTAGCC 41103
QY 216 lyArgAsnAspSerGlnTrpArgProGln----- 225
Db 41104 GGCAATATGATGAAGGAGGAGGAGCCATATCCTAGGCCACTTGTTCAGGTGAGATTTTCAC 41163
QY 225 ----- 225
Db 41164 AGTTGATTTGATCGCTGTACTACTAACTTTCTCTGTGTTGGCTATCTCTGAACATAG 41223
QY 226 -----GlnHisLeuValHisAlaThr----- 232
Db 41224 CCTATCCTAGAACTTTCTATTGCGATGTTAT-ATTGAAAGAGAGCTAGTAGTATTTCTCTC 41282
QY 233 -----ProTyrGluSerLeuArgAspI 240
Db 41283 TTGCTATTAAATGTTTATATATATGCGCTATCAGGTGGACCTATGATAATCTGAAGAGC 41342
QY 240 leAlaValIysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerS 260
Db 41343 TTGCCCTGAAATTTTCCAAAACAAGTGGCAGCTGTTCCGTTATATATTTCTTTCTTGC 41402
QY 260 erAspGlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuIysCys- 279
Db 41403 AGGATGTTTCATATCCGAGTTACTGTCATCTTCTTCCGCTATCCAGGCATATTAAATGTA 41462
QY 279 ----- 279
Db 41463 AGGTCCCTTCATTGTGTAATAGACACTGATAAGCGGCCAATTTGGCAITTCCTCACTTT 41522
QY 279 ----- 279
Db 41523 TTAATTCATGAGGTGCATATGATTTCTGCTATTTCCAAAATTTTCCCACTTCTTCAITTTCT 41582
QY 279 ----- 279
Db 41583 CCAANTTCTGGAAGGATGGGATCTTTAGGAGGAGGGAATTAAGAACTAAAGTTGCA 41642
QY 279 ----- 279
Db 41643 ATTGACTTTACCAACTTTTACACAGTTTTCCTGTTTAAATAATTCCTTTTCGAAATTTAT 41702
QY 280 -----IleCysArgTyrP 284
Db 41703 GCCTTTATCAGACACTTAAGCTCTAAATTAAGCATAATCTTCCAGGTATATGSCAGATACT 41762
QY 284 heLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuG 304
Db 41763 TTAGACATTTGCTAGCTCTTTACTATCTCTTCCAGCAACCACTTGTTCATTTCCCTCGG 41822
QY 304 lySerTrpValProIyIleIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProA 324
Db 41823 GTACGTGGTCCCTAGAAATCGGAAATCAAGTAGCAAACTCTCCGCTACATTGAGACCTC 41882
QY 324 snAlaSerLeuSerSerAlaLeuAsnMetLeuValGln----- 336
Db 41883 ACGCCTCTTTGGTTCGCGCTCGCATTTATAGTTCAGAGGTAATCCCTTTTCTTATCTTG 41942
QY 336 ----- 336
Db 41943 TTGCAGCAACATCTTATTCAAAGCTCATCTGCACTCAGACTATCAACCTCTACATTACA 42002
QY 336 ----- 336
Db 42003 ATCAGATTTTCAGTTTAAATATTTTGTGTAGATGGTAATATAATCAAGTAGCTGAACAC 42062
QY 336 ----- 336
Db 42063 ATGGGTATCTCAAGTCAATGCAAGCAACAAACCAACTTTAGATGTTAAGCAACCACTTTT

QY 336 ----- 336
Db 42123 ATTATTTTATACAGATGTAATTGCTGCTCAACATGAGTAACATAACGAGGTTTAAAAACAT 42182
QY 336 ----- 336
Db 42183 AGAAACATTTTGTCAATCTTAAAGCAGATATATTTATAGAAATTTCTCGTGATCTTTTCC 42242
QY 336 ----- 336
Db 42243 CTGCTTCCTCATAGATGAAATGAATGAATGAGTCATGCTTAATTTATGTTTCTCGGTTA 42302
QY 337 -----AlaGlyValSerSerIleProIleValAspAspAsnAspSerLeuLeuAspThr 355
Db 42303 TTACAGCTGAAGTCRGTTCAATTCCTCGTAGTGATGACACGACTCGCTTATTGACATAT 42362
QY 355 yrSerArg----- 357
Db 42363 ACTCTCGAAGGTAAACCTCTATCCATCAATCGACATGATGCTTATGCAITTCCTCTTTG 42422
QY 358 -----SerAspIleThrAlaLeuAlaLysA 366
Db 42423 ACGTCTTGTGTTCTGTGATATTTGTAATTTGGCCAGTGATATACTCTCTGGCTAAG 42482
QY 366 spLysValTyrThrHisValArgLeuAspGluMetThrIleHis----- 380
Db 42483 ACAAGGCATAGCACAGATTTCATCTTGTATGACATGACAGTTCACGAGTAAACCTCTCAT 42542
QY 380 ----- 380
Db 42543 CAGAAACCAAGATCCATCATCATGCTCATCAAAATTACGAGTTTACTTAAATGTAACCTG 42602
QY 380 ----- 380
Db 42603 CCTATCATCATGAAATTTGCTGTTTCTGATTTATTTGAAACATACATCTTTTACAGAA 42662
QY 381 -----GlnA 382
Db 42663 TGAAGAAGAGATGAAACGTTTATATAAAGTGGAGGATAATGTGTTGTGGTTCAGG 42722
QY 382 laLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysG 402
Db 42723 CGCTACAGTTTGGGCAAGATGCGAGCCGCTTATGGAATCTTCAACGGGAGAGATGTC 42782
QY 402 lnMetCysLeuArgSerAspProLeuLeuIysValMetGluArgLeuAlaAsnProGlyV 422
Db 42783 ACATGTCTTGGCTCAGACTCTCTTGTGAAGATGATGAGCGGTGGCGAATCCAGTA 42842
QY 422 al----- 422
Db 42843 C-ACACTATCTCTATATACAATCTTCAGCTTTGTTGTTGTCACACAGGAACATCAGATGA 42901
QY 423 -----ArgArgValPheI 427
Db 42902 GATGTAACCTGGGTTTGGGTTTGTGTATATAAACAACAACAGGTGTAAAGGAGCTGGTA 42961
QY 427 leValGluAlaGlySerIysArgValGluGlyIleIleSerLeuSerAspIlePheTysp 447
Db 42962 TAGTGAAGCAGGAGGCAACGCTGTGAAAGATATCAITTCATCTGAGCGATGTTTCCAAT 43021
QY 447 heLeuLeuSerLeu 451
Db 43022 TCCTGCTCGGTCTT 43035
RESULT 14
AK121122
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023074117, full
insert sequence.
ACCESSION AK121122
VERSION AK121122.1 GI:37990745
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM	Oryza sativa (japonica cultivar-group)		REFERENCE	Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Ikeda, R., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, A., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, W., Waki, K., Yasunishi, A. and Hayashizaki, Y.	
	The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, A., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sakai, D., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.			Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice	
TITLE	japonica rice		REFERENCE	US-09-857-525C-2 (1-451) x AK121122 (1-1923)	
	Science 301 (5631), 376-379 (2003)			QY 10 lIeGluGly-CysPro-----ThrValPheGlnAlaIleCysSerLeuSe 24	
JOURNAL	2752273		AUTHORS	455 GTGGAGGAGGTGCCCCATGGCGGCTCGTGGGCGCGAGTTCAGGTGTTCGATCTGCC 514	
	12869764			QY 24 rProGlyIleHisGluTyrLysPhePheValAspGlyGlyTyrArgHisAspGluA-GGL 44	
REFERENCE	2		AUTHORS	515 CCCCAGGGGTTTATCATGATACCCGGTTTTGGTTGATGGTGTCTGGCGATGTGATGAGAGAA 574	
				QY 44 nProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyrAs 64	
AUTHORS			AUTHORS	575 ACCTTGTGTACGGGATGAATGGTTGATTAGCAATCAAGATCTTCGTGGATATATACACA 634	
				QY 64 nGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspValAs 84	
TITLE			REFERENCE	635 TCCGGTT-----GTACA 646	
				QY 84 pAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrVal---SerGluGlyTh 103	
JOURNAL	3 (bases 1 to 1923)		AUTHORS	647 GCCAGAACTTCCATCAGATAGTAGAGTGTGATGGAGGTACTATATCTGACACAGTATG 706	
	Direct Submission			QY 103 rLeuArgValSer----- 107	
JOURNAL	Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp]		REFERENCE	707 TATCAAGTTTCATTGTGTAATTTTGTGTAATTTTATGCTACTAATAACTTATAGTC 766	
	This clone is one of the 32K full-length cDNA clones from japonica rice.			QY 107 ----- 107	
JOURNAL	http://cdna01.dna.affrc.go.jp/cDNA/		AUTHORS	767 TGTATTGCAATTATTCAAATTATCTCAAGCAGATGCCGCCAGATCAGTTATCTCAAAC 826	
	NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M. and Nakahama, Y.			QY 108 -GluAlaAla-IleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrC 127	
JOURNAL	FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,		REFERENCE	827 CAGGGCGTCAGATAGCAATTTTCGCCCATCGAGTCTCTGAGATATATTACATAATACCA 886	
				QY 127 yfTyrAspLeuLeuProAspSerGlyIlysValIleAaLeuAspIleAsnLeuProVal 147	
JOURNAL			AUTHORS	887 TATATGATGTTGTTCTCTGTTCTTCTAGCAAGATAGGAGGTTTGGATGCTCTTCTCTCA 946	
				QY 147 yfGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTrpAspSerP 167	

```

Db 947 ACCAAGCCTTTAAATCATGCGATGATGAGGCTCTTCTTCTTGGTCTCTCTTGGGATGATC 1006
Qy 167 heArgelyGlnPheValGlyLeuLeuSerProLeuAspPheLeuLeuLeuLeuArgGlnL 187
Db 1007 AGCAGCAACCGTAAACAGGATGCTAACCGCATCCGATTTGTATTAATTTTGAGAAAGC 1066
Qy 187 euGluThrHisGlySerAsnLeuThrGluGlnLeuGlnThrHisThrIleSerAlaT 207
Db 1067 TGCAGAGAAACATTCGAACCTCTTGGCATGAGGAACTTGAATGCAATTCAGTCTGCTT 1126
Qy 207 rplysGluAlaValArgGlnThrAsnGlyArgAsnAspSerGlnTTPArgProGlnSLNH 227
Db 1127 GGAAGAGCAAAATTCAGTCTTATGCGGAGCTGATGCTGCGCTCCATTCAGAGAGGC 1186
Qy 227 isLeuValHisAlaThrProTyGlySerLeuArgAspIleAlaValLysLeuGlnA 247
Db 1187 CATTAATCATGTTAAGGATTCAGATAATTAAGGATGTCGCAATGGCTATAATTAGAA 1246
Qy 247 snGlyIleSerThrValProValIleTySerSerSerSerAspGlySerPheProGlnL 267
Db 1247 ATGAAATATCTTCAGTCTCT--ATCTTTAAGCCCTCAACGGATTCATCAGGGATGCTT 1303
Qy 267 euLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyPheLysAsnS 287
Db 1304 TGTCTGCTTGTCAACCTTCAGGATGTTGAAATTTATTGCTCAAAGCTCAAGAAC 1363
Qy 287 erThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpV 307
Db 1364 AGCCTGAAGGTTACTATTCTGCAAAACCCAGATTTGTCAGTATGCTTATGTCATGCT 1423
Qy 307 aProIlyIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerL 327
Db 1424 CACCATACTGCAAGGCAAGCAATAGACAGCTTAGAACTTCGCGACCAAGCACTCCTC 1483
Qy 327 euSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspA 347
Db 1484 TAAATTCATGCTTGATTTGCTTGAAGATAGATAGTAAAGTCAATTCCTATATGTCAGC 1543
Qy 347 spAsnAspSerLeuLeuAspThrTySerArgSerAspIleThrAlaLeuAlaLysAspL 367
Db 1544 ATAATGGCGCTCTCTTGTATGCTTACTCGCTCAGTGATATCATGCTCTAGCGCAAGATG 1603
Qy 367 ysValTyThrHisValArgLeuAspGlnMetThrIleHisGlnAlaLeuGlnLeuGlyG 387
Db 1604 ATGCTACATCGTATTGAGCTTGAACAGTGAGCGTGAGCATGCTTCGAGCTGCAAT 1663
Qy 387 InAspAlaAsnThrProPheGlyPheAsnGlyGlnArg--CysGlnMetCysLeuA 406
Db 1664 ACCAGGTG-----AATGCCGCAAGACACTGTGCATACCTGCTTGA 1702
Qy 406 r9SerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgValP 426
Db 1703 GCACGTAGTACCTTCCTGGAGTTTGGAGCAATTTGTCAGCTCCAGGGGTGGCGAGTGG 1762
Qy 426 heIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheL 446
Db 1763 TCGTATTGACCAAGAGCGAGATTGTGCAAGGAATATCTCATTTGAGGAGCAACCATTTA 1822
Qy 446 ysPheLeuLeu 449
Db 1823 CATTTCTCATT 1833

```

```

RESULT 15
AC146571
LOCUS
DEFINITION Medicago truncatula clone mth2-104a2, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC146571
VERSION AC146571.6 GI:39752723
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 139089)
Lin.S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Medicago truncatula BAC Clone mth2-104a2
Unpublished
2 (bases 1 to 139089)
Lin.S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Direct Submission
Submitted (30-AUG-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 139089)
Lin.S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Direct Submission
Submitted (12-DEC-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Dec 12, 2003 this sequence version replaced gi:38488825.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 6334: contig of 6334 bp in length
* 6434: gap of unknown length
* 6435: contig of 18620 bp in length
* 25054: contig of unknown length
* 25154: gap of unknown length
* 25155: contig of 47115 bp in length
* 72269: contig of unknown length
* 72370: gap of unknown length
* 72370: 139089: contig of 66720 bp in length.
*
* Location/Qualifiers
* 1..139089
* /organism="Medicago truncatula"
* /mol_type="genomic DNA"
* /db_xref="taxon:3880"
* /clones="mth2-104a2"
* /clone_lib="Medicago truncatula BAC library H2"

```

```

ORIGIN
Alignment Scores: 2e-46 Length: 139089
Pred. NO.: 682.00 Matches: 236
Score: 26.60% Conservative: 46
Percent Similarity: 22.26% Mismatches: 65
Best Local Similarity: 29.12% Indels: 715
Query Match: 2 Gaps: 11
DB:
US-09-857-525C-2 (1-451) x AC146571 (1-139089)

```

```

Qy 101 GluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGlu 120
Db 64274 GAACTAATCAAGATGTCGCCATCTGATCTGGAGGTCTCTCGCGCGGTATATCTAAA 64333
Qy 121 TyrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLys----- 136
Db 64334 TTCTTGTCGAACATACTGCATATGATTTGCTTCTCTGAATCTGGAGGTTTGAATGATT 64393
Qy 136 ----- 136
Db 64394 TTTCATTATCAATTTTGTCTGCTGCTCATTCATTAATTAATTAAGACTTATTCATGTTG 64453

```

QY	137	-----ValileAlaLeuAspIleAsn	143	QY	230	-----	230
Db	64454	GCTTATTATCGCTGTGTATATAATATAATAATTCAGGTCTATGCCCTTGGATGTAAT	64513	Db	65534	TGTCAGTACCAGTCCCTCTCTGTATTTGTTATAGTTTTTTTACAGCAGCATTCATG	65593
QY	144	LeuProValIysGlnSerPheHisIleLeuHisGluGln	156	QY	230	-----	230
Db	64514	TTACCAAGTTAAGCAAGCATTCATCTTCTTTATGACAGGTATTTCCGCCCTTTTAAAA	64573	Db	65594	TGGTTTTGTCCCATCAGTTATGTCANATTTCTGCTTCATTCCTTCTCAATATATCATCAT	65653
QY	156	-----	156	QY	230	-----	230
Db	64574	TATATCTTCTTATTTCCATCAGGCGACAGCCATGCCCGTATATGTAATTAATTTGT	64633	Db	65654	TCTTTTCTCCTACACTATTTATCAATTTGTATCATATAAAATTTTGATAGCTTTTTTTG	65713
QY	157	-----Gly-IleProValAlaProLeuIlePheSerPheArgGlyGlnPheValGln	173	QY	230	-----	230
Db	64634	TTTCTTTCAGGATGATCTATGCGCAGCTATGGGATTTTTCAGAGAGTCAATTTGTGTG	64693	Db	65714	CATCAATGAATGATGTTTAAACATTCATGTTGACCAATGCAATTTTAAATGTGTAC	65773
QY	173	LeuLeuSerProLeuAspPheIleLeuIleLeuArgGlu	186	QY	230	-----	230
Db	64694	AGTGCTTAGTGAATGGACTTCATCTCTCATTTGAAAGAGGTTTGTGGTAATAAAACTA	64753	Db	65774	TTAATTCGTATGATGTTTGTATTTCCCTTCTATCAAAAAGACCGGAGTAACTCTTGAA	65833
QY	186	-----	186	QY	231	-----AlaThrProTyrGluSerLeuAr	238
Db	64754	TTAAGCAACCTTCTGTGCCAGTTTATAGTAAAGTCACCTACAACTATTTTATCG	64813	Db	65834	CTTTTATGATACTTTTGTGTTGCTTGTGATCTTCAGSCTGGGCCCAAGGAATGCCATAA	65893
QY	186	-----	186	QY	238	gAspIleAlaValIysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSe	258
Db	64814	GGGTGACGTTTAAAGCTATTAAACAGCTGCTTACAAAATGGTAAAAAATAAAAACTTAA	64873	Db	65894	AGAGTGGCTTTGAAGTTTTTGCAAAACAAGGTGTCACTGTCTCTATCATTTTC	65947
QY	186	-----	186	QY	258	rSerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLy	278
Db	64874	GGAGTAAGAAATGACATTAATAATTTCAAGAGAAATCATGCTATTCATGACATTAATAATTT	64933	Db	65948	TTTATGAGGATGGTTCAITTTCCCAACTGCTGCATCTTGTCTCCCTATCAGGAATTTTAA	66007
QY	186	-----	186	QY	278	sCys-	273
Db	64934	TAAAGAAATCATGTTCTTATATTTGGGCTAATAATATAGATCATATCGGAGCTGGT	64993	Db	66008	ATGTAAGCTTCTAGTCTGTGTGTCATATAGATAAATGGTTTGGAACTCTGAATGATGCT	66067
QY	186	-----	186	QY	280	-----IleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleL	294
Db	64994	ATGAGATTTATTGAGCTCATACAATAATATTTGTAATATCATTTTGGGTGAAGGCTT	65053	Db	66068	GGATAAATTTTTCAGGTATATGTCAGGCATTTTGAGCACTCTGCTGGTCTCTGGCCATTC	66127
QY	186	-----	186	QY	294	euAsnGlnProValCysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuA	314
Db	65054	TAACATATGTCATGTTTAACTTTTATGCTGCTGTTGGCTTAAATGCTCTTGAGT	65113	Db	66128	TTCAACTTCGATGTTCTGATACCTTTTGGTATACGGTGCTTAAAGCTGGGATCCAA	66187
QY	186	-----	186	QY	314	snSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetL	334
Db	65114	TCCTTACATCTAGATATGTTGTTGCTGCTATATTTCTCTGTAACATACTGCCATC	65173	Db	66188	ATGACACGCCACTGATAAGGCTAAGGCCAAATGCTTCTTTGGTGGATGCCCTGTGATGT	66247
QY	186	-----	186	QY	334	euValGln	336
Db	65174	ATGATTTTCATTCATACCTTTAGTATCTATGATTCATCAATGCCCTACTCTTCTAGTGAAT	65233	Db	66248	TGTTTCAAGGTGATCATCTTATGTTTAAAGTTTGTCTATTTTCTTCTGTTTGTTC	66307
QY	186	-----	186	QY	336	-----	336
Db	65234	GCATGATGTTTTTGTGTTGTTGTTAAATTTAATTAATTTGACATGCTGTAATTTCA	65293	Db	66308	TGATTTATGTCCTTTTCCCATCATGATATTTTCTGAAAGGTGTCCCAAGGGGATCTTTTGA	66367
QY	187	-LeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerAl	206	QY	336	-----	336
Db	65294	GTTAGGGACTCATGGATCATATTTGACTGAAGAACAACCTTGAGACTCATACTATAGCAGC	65353	Db	66368	AAACTCTTGTAAATTTTAAAAAAGTTATCCATTAATTAAGGGGAGGTTACTGTGTATA	66427
QY	206	atpLysGluAlaIlyArgGlnThrAsnGlyArgAspSerGlnTrpArgProGlnG	226	QY	336	-----	336
Db	65354	CTGGAAGAGGAAATCAAAGCAACGCGGCACCTTGATAAATAAGGATCAAAATCC	65413	Db	66428	AATAGTGCACCATCATTTCTGTATGGACCAATAGAAACATTCAGTTTATAGGACAAAT	66487
QY	226	nHis--LeuValHis	230	QY	336	-----	336
Db	65414	TCATTTGTTTGTTCATGTAAGTCTCTGATGACTGCGATAACTTTTTCATGATCTAAG	65473	Db	66488	GATTCGTGCAACTCAAGCTAAGAAATGTTCATTTAACTGTAAGTTTAAACAGCAGAAA	66547
QY	230	-----	230	QY	336	-----	336
Db	65474	TGTTCTCTCAGTTTCATAGTTTGTAGTTTCTCTACTTTTTCATGTTTATAGACTTCAGTT	65533	Db	66548	TGTACTCCATGTTGAAATTAGAATTTTAAAAATTTAAATGTTCATTTTACAAATACATAT	66607
				QY	336	-----	336

```
Db 66608 ATTTCGGTTGTTATTTTATTGGTCTCAGGAGGGACAAATATTCGATGAATAGTTTTC 66667
QY 336 ----- 336
Db 66668 TCTTGATTAGTTTGTGTAGTAAAGACACAGTCTTGTAAATACACGGTACCTTTAACCTG 66727
QY 337 ----- -AlaGlyVal 339
Db 66728 TTTTTCGGTCTAGCTTTTGATATCCCTATGATGAAATGCTCTTCTATTGTCAGCTAAAGTT 66787
QY 340 SerSerIleProIleValAspAsnAspSerLeuLeuAspThrIleValSerArg ----- 357
Db 66788 AGCTCAATACCAATTGTGGATGAATGATTCATTCCTTGACATATATTCAGAGAGTAG 66847
QY 357 ----- 357
Db 66848 TATTTTTCCTCCCCCAATTTCCTTTTGTGTGTGANTATGATATATTATTGAAAGAT 66907
QY 358 ----- SerAspIleThrAla 362
Db 66908 CCTGAGCATATATGAATATGATGTGCTGCTATTATTTCTTTGTCAGTGATATTACTGCA 66967
QY 363 LeuAlaLysAspLysValIleThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
Db 66968 TTGGCTAAGATAAAGCATATGCTCGGATATCCCTGATGAACAACTAATATTCAATCAGGT- 67026
QY 382 ----- 382
Db 67027 ACCTACTATGATGACATTTTGAGTTATAATAATGCAATACATTTGTTGTTTGTGTTG 67086
QY 383 ----- -LeuGln 384
Db 67087 TTTGTTAGCAAAATTAAGTGAATCTAAATATTTGCACTACATGGTTGTAGGCATTAAAT 67146
QY 385 LeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGlnMetCys 404
Db 67147 TTGGACAAGATGCAAAATTCCTTATGCTTTACAAATGGGCACAGATGTCATATGCT 67206
QY 405 LeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnPro ----- 420
Db 67207 TTGAGGTCGTGATTCATTGCACAAAGTGATGGAGCGGTGGCTAAACCTGGTATTACTTTG 67265
QY 420 ----- 420
Db 67267 TTTACTGATATAATAGCATAGCTCCCTCCACTTCTTAAATGGAAGGTAGTTCCTT 67326
QY 421 ----- GlyValArgArgValPheIleValGluAlaGly 431
Db 67327 TATTCATCGGTTTTCGCACCTTGCAGGGGTT-AGGAGACTCGTGATTTGGAGCTGCT 67385
QY 432 SerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPheLeuLeu 449
Db 67386 ACCAAGCGTGTGGAGGGATTATTCATTAGTGATGTTTTCAGATTCTTATTG 67439
```

Search completed: July 9, 2004, 10:15:08
Job time : 5816 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2004, 05:05:39 ; Search time 609 Seconds

(without alignments)
3146.038 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREHLPMSIECPVFWQAI.....SKRVEGIIISLDIEKFLLSL 451

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09857525/runat_07072004_161017_283/app_query.fasta_1.647
-DB=N Geneseq 25Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -CUTPMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09857525_@CGN_1_1_470_@runat_07072004_161017_283 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	1576	3	Aaa51347 Z. mays s
2	2085	89.0	2149	3	Aaa51348 Rice suc
3	2031	86.7	1632	3	Aaa51353 Wheat suc
4	1581	67.5	2538	3	Aaa51351 Soybean s
5	1525	65.1	2163	3	Aaa51350 Soybean s
6	923	39.4	538	3	Aaa51354 Maize suc
7	530	22.6	514	3	Aaa51357 Wheat suc
8	460	19.6	282	6	ABL70761 Corn tass

9	392	16.7	702	3	AAA51349	Aaa51349 Rice suc
10	373.5	15.9	1167	6	AAU48614	Aal48614 Human ins
11	373	15.9	1435	2	AA06882	Aax06882 Disease a
12	373	15.9	2223	4	AHL14839	Ahl14839 Human cDN
13	373	15.9	2223	6	AAU48615	Aal48615 Human ins
14	367	15.7	1578	6	AAU48616	Aal48616 Human ins
15	367	15.7	1578	6	ABK84324	Abk84324 Human cDN
16	367	15.7	1578	7	ACA56515	Acas6515 Human sig
17	367	15.7	1691	7	ACA98774	Aca98774 Human pan
18	364	15.5	1576	2	AAT85927	Aat85927 Mammalian
19	361.5	15.4	1905	6	ABS51409	Ads51409 cDNA enco
20	361.5	15.4	3261	4	ABL18857	AbL18857 Drosophil
21	340	14.5	1647	5	AAH43685	Aah43685 PRKAG3 CD
22	340	14.5	2109	4	AA03296	Aad03296 Human AMP
23	340	14.5	2115	4	AA03320	Aad03320 Human AMP
24	340	14.5	2314	6	ABQ61173	Abq61173 AMP kinas
25	340	14.5	2781	7	AAU49456	Aad49456 Human kin
26	336	14.3	1774	6	ABL39755	AbL39755 Human NS
27	335.5	14.3	1873	6	AAU36460	Aad36460 Pig PRKAG
28	335.5	14.3	1873	6	AAU36459	Aad36459 Pig PRKAG
29	334.5	14.3	1857	4	AA03295	Aad03295 Pig AMPK
30	334.5	14.3	1873	4	AA03319	Aad03319 Pig AMPK
31	334.5	14.3	1873	6	AAU36458	Aad36458 Pig PRKAG
32	334.5	14.3	1873	6	AAU36457	Aad36457 Pig PRKAG
33	334.5	14.3	1873	6	AAU36456	Aad36456 Pig wild-
34	334.5	14.3	2022	4	AA03321	Aad03321 Sus scrof
35	309	13.2	39651	4	ABL18856	AbL18856 Drosophil
36	303.5	13.0	969	7	ACC60984	Acc60984 Gene sequ
37	302	12.9	1467	5	AAS84265	As84265 DNA encod
38	302	12.9	1467	9	ADE07539	Ade07539 Novel cod
39	302	12.9	1467	9	ADE09727	Ade09727 Novel DNA
40	269.5	11.5	1266	3	AAA51352	Aaa51352 Soybean s
41	226.5	9.7	542	3	AAA51355	Aaa51355 Rice suc
42	223.5	9.5	178	6	ABL75404	AbL75404 Corn tass
43	211.5	9.0	735	4	AAU07561	Aah07561 Human cDN
44	184.5	7.9	2303	5	AAS84267	As84267 DNA encod
45	183	7.8	765	6	ABL89686	AbL89686 Human pol

ALIGNMENTS

RESULT 1

AAA51347

ID AAA51347 standard; cDNA; 1576 BP.

XX

AC AAA51347;

XX

DT 26-SEP-2000 (first entry)

XX

DE Z. mays sucrose non-fermenting 4 protein cDNA from clone cskic.pk001.c15.

XX

KW Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

KW carbon catabolite repression; development; nitrogen partitioning; ss.

XX

OS Zea mays.

XX

PH Key

CDS

FT

FT

FT

FT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX WPI: 2000-431594/37.
 DR P-PSDB; AAY96783.
 XX
 ET New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 XX and development.
 XX
 PS Claim 2; Page 31; 48pp; English.

XX This cDNA encodes maize (Zea mays) sucrose non-fermenting 4 (SNF4)
 CC protein which is involved in carbon catabolite repression in plants and
 CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
 CC cskic.pk31.c15 prepared from corn unpollinated developing silk 24 hours
 CC after emergence. The polynucleotides are used in plants to control carbon
 CC and nitrogen partitioning pathways during plant growth and development.
 CC The catabolite repression proteins would facilitate studies for better
 CC understanding the mechanism of catabolite repression in plants and could
 CC provide genetic tools to enhance or otherwise alter the accumulation of
 CC carbohydrates, lipids and proteins during plant growth and development
 XX

SQ Sequence 1576 BP; 418 A; 336 C; 362 G; 460 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,45e-223 Length: 1576
 Score: 2342.00 Matches: 451
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DE: 3 Gaps: 0

US-09-857-525C-2 (1-451) x AAA51347 (1-1576)

Qy 1 ThrArgGluHisLeuProMetSerProLeuGluGlyCysProThrValPheGlnAlaIle 20
 Db 3 ACGAGGGAACATTTGGCCGATGCTCTCTATCGAAGGCTGCCCATCTGTTTTCAGGCTATT 62
 Qy 21 CysSerLeuSerProGlyIleHisGluCysPheValAspGlyGluTrpArgHis 40
 Db 63 TGCAGTCTCTCTCTGGGATTCACAGGACAAATTCITTTGTGGAGGGGAATGGGGCAT 122
 Qy 41 AspGluArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTrpLeuThr 60
 Db 123 GATGAGCGTCAACCTACCATATCTGGGAGTTTGGCATAGTTAACAACACTTACTTGACA 182
 Qy 61 ArgGluTrpAnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsn 80
 Db 183 AGCGAATATAACCAATAAACCCTTATCAAGTCCAGCACACCTGGGAAGCAGGATGAAC 242
 Qy 81 MetAspValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSer 100
 Db 243 ATGGATGTGGATATGAAATTTTCAAGTACGTTAGCTTGTGATGGCACCCTTTCA 302
 Qy 101 GluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGlu 120
 Db 303 GAAGGTACTCTGAGAGTTTCAGAGGCTCAATAACAATATCTAGTGTGCGGTTTCGAA 362
 Qy 121 TyrLeuAsnLeuHisThrCysTrpAspLeuLeuProAspSerGlyIleValIleAlaLeu 140
 Db 363 TATCTGAATTTGCATACATGCTATGATTACTCCAGATCTTGGCAAGGTTATTGCCCTA 422
 Qy 141 AspIleAsnLeuProValIleGlnSerPheHisIleLeuHisGluGlnGlyIleProVal 160
 Db 423 GACATTATTTACTGTGAGCAATCATCTCATATTCTCCATGACAGGGGATTTCTTGTA 482
 Qy 161 AlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPhe 180
 Db 483 GCTCCTCTCTGGGACTCATTCAGAGGTCAATTTGTGTGCTCTCTAGCCCATTTGGATTTC 542
 Qy 181 IleLeuIleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGlu 200

Db 543 ATACTCATATTGGGGAGCTAGAACTCATGGCTCGAACTTGACAGAGAGCGACTGAA 602
 Qy 201 ThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSer 220
 Db 603 ACACACACTATATCTGCTGGAAGAGGCTAAGCGGCAAACTAATGGAAGAATGATAGT 662
 Qy 221 GlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGlnSerLeuArgAspIle 240
 Db 663 CAGTGGCGACCGCAACAGCATCTAGTGCATGCCCTTATGAGTCTCTTGAGGGACATT 722
 Qy 241 AlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTrpSerSerSerSer 260
 Db 723 CGAGTAAGCTTTTCGAAATGCGATTCTACAGTGCCAGTTATTTTATCATCATCTCA 782
 Qy 261 AspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIle 280
 Db 783 GATGATCATTTCCCGCAGTTATTGCATCTTGCATCATTCTGGAATTTTGAATGTATT 842
 Qy 281 CysArgTrpPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSer 300
 Db 843 TGTAGATACCTTCAAAACACTCAACTGGTAATTTGGCTTATTTCTGAACCAACAGTGTCTCC 902
 Qy 301 IleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMet 320
 Db 903 ATTCCGCTGGGTTCCTGGGTCCGAAATTTGGTGATCTGAACAGTCTGCTCATTTGGCTATG 962
 Qy 321 LeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSer 340
 Db 963 TTGGACCTTAATGCTCTAGCTCTGCCCTTAACATGTTGGTTCAGAGCTGGAGTAAGC 1022
 Qy 341 SerIleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIle 360
 Db 1023 TCATPACCAATTTGTGATGACAAACGACTCCCTGCTTGACACTTACTCTAGAAGTGACATC 1082
 Qy 361 ThrAlaLeuAlaLysAspLysValTrpThrHisValArgLeuAspGluMetThrIleHis 380
 Db 1083 ACAGCCCTAGCAAAAGACAAGGCTACACATGTTGCGCTGGATGAGATGACCATTCAC 1142
 Qy 381 GlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArg 400
 Db 1143 CAGGCTTTTGCAGCTTGGACAAGATGCCAATACACCTTTTGGATTCTTTAAACGCCAGAGA 1202
 Qy 401 CysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnPro 420
 Db 1203 TGCCAGATGTGCTCCGCTCGATCTGATCTTGTCTGAGGTGATGGAGCGACTGGCTAATCT 1262
 Qy 421 GlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSer 440
 Db 1263 GGGGTGCGCGGGTGTTCATTGTGGAAGCTGGAGCAAAACGTTGGAGGCAATCATATCA 1322
 Qy 441 LeuSerAspIlePheLysPheLeuLeuSerLeu 451
 Db 1323 CTAAAGTGAATTTCAAGTTCTTGTGAGCTTG 1355

RESULT 2

AAA51348

ID AAA51348 standard; cDNA; 2149 BP.

XX AAA51348;

AC AAA51348;

XX 26-SEP-2000 (first entry)

DE Rice sucrose non-fermenting 4 protein cDNA from clone r10n.pk083.i18.

KW Rice; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX carbon catabolite repression; development; nitrogen partitioning; ss.

OS Oryza sativa.

XX Key Location/Qualifiers
 XX CDS 185..1666
 FT /*tag= a

/product= "Sucrose_non-fermenting_4_protein";

WO200036116-A2.

22-JUN-2000.

15-DEC-1999; 99NO-US029825.

16-DEC-1999; 98US-0112564P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Allen SM, Heppard EP, Xiao G, Weng Z;

WPI; 2000-431594/37.

2-PSDB; AA96784.

New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins involved in carbon catabolite repression in plants and seeds, useful for controlling carbon and nitrogen partitioning pathways during plant growth and development.

Claim 2; Page 33-34; 48pp; English.

This cDNA encodes rice (*Oryza sativa*) sucrose non-fermenting 4 (SNF4) protein which is involved in carbon catabolite repression in plants and seeds. The cDNA was isolated, based on similarity to SNF4 proteins from *Arabidopsis thaliana* and *Saccharomyces cerevisiae*, from clone r10n.pk083.i18 prepared from rice 15 day old leaf. The polynucleotides are used in plants to control carbon and nitrogen partitioning pathways during plant growth and development. The catabolite repression proteins would facilitate studies for better understanding the mechanism of catabolite repression in plants and could provide genetic tools to enhance or otherwise alter the accumulation of carbohydrates, lipids and proteins during plant growth and development.

SQ Sequence 2149 BP; 523 A; 492 C; 501 G; 633 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.66e-197 Length: 2149
Score: 2085.00 Matches: 392
Percent Similarity: 94.87% Conservative: 34
Best Local Similarity: 87.28% Mismatches: 23
Query Match: 89.03% Indels: 0
DB: 3 Gaps: 0

US-09-857-525C-2 (1-451) x AAA51348 (1-2149)

QY	3	GluHisLeuProMetSerProGlyGluGlyCysProThrValPheGlnAlaIleCysSer	22
DB	320	GAACATTGGCAATGTCCTCTGTGAAGGTGTCCACCGCTTTTCAGGCCATATGCAGT	379
QY	23	LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu	42
DB	380	TTGTCCCTCGGATTCATCATCAAGATTTTGTGTGGATGGGAATGGCGCAGATGAG	439
QY	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu	62
DB	440	CGCAACCTTACCATAACAGAGAGACTATGGTGTGTGTAACTTTATCTGCTGACTAGGAC	499
QY	63	TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp	82
DB	500	TTTGACACAGATAAATAAATATTAGGCCCTAGTACACCTGGAGCTAGGATGAACATGGAT	559
QY	83	ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
DB	560	GTGGACAAACGACAAATTTCAACAGTACTCTTTGTCTGTATGTCATTATTTCAGGAAGT	619
QY	103	ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu	122
DB	620	CCTCAGAGAATTTTCAGAGGAGCTATACAGATCTCTAGGTGCTGTAGCGGATTTCTG	679
QY	123	AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle	142

DB	680	AAATGGACAAACTGGGTATGATTACTCCAGATTCTGGCAAGGTCTATTGCTCTAGACGT	739
QY	143	AsnLeuProValLysGlnSerPheHisIleLeuHisGlnGlyIleProValAlaPro	162
DB	740	AAATTTGCTGTGAGGCAATCTTTTCATATTCTTCATGAACAGGAAATTCCTGTGGACCT	799
QY	163	LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu	182
DB	800	CTATGGGATTCATTGAGGGCCAGTTTGTGGCTTTTGAGCCCTGGATTTTATATCTC	859
QY	183	IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis	202
DB	860	ATATTGAGAGAGCTGGAACCTCATGGCTCCAAATCTGACAGAGAGAGCTTTGAACACAT	919
QY	203	ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGluTyr	222
DB	920	ACTATATCTGCGTGGAGGAGGCCAAGCGTCAAACTTATGCCAGAACGAGGGTCTCTGG	979
QY	223	ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal	242
DB	980	AGGGCAAAATCACCATTAGTTTCATGCCACCCCTTATGAATCCCTGAGGGAAATGCTATG	1039
QY	243	LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly	262
DB	1040	AGATATCTGCAAAACGGTGTCTTCTACCGTTCCAAATATGTTTCTCTCATCACCAGATGGC	1099
QY	263	SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysArg	282
DB	1100	TCATATCCCAATGTTGCACTGTCTCCCTTCTGGAAATTTTGAATGTTATTTGAGA	1159
QY	283	TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro	302
DB	1160	TATTTTAAATTTCTCAGGTATTTTACCTATTGTGAGCCACCTGTATGCACATTCCT	1219
QY	303	LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg	322
DB	1220	CTGGTACCTGGGTTCAAAAATTTGGTGATCTTAATGGCGCTCCATTGGCTATGTCGG	1279
QY	323	ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle	342
DB	1280	CCTAACAACATCTCTTACGCGTGCCTGAAATTTGTGTGTTCAAGCTGTGTGAGTCAATA	1339
QY	343	ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla	362
DB	1340	CCAATTGTGATGCAATGACTCCTCTTGACACATATTCAGAAAGTGACATCACAGCT	1399
QY	363	LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla	382
DB	1400	CTTGCAAAAGACAGGTGTATACACACATTCGCTAGATGAGATGACCATTCATCAGGCT	1459
QY	383	LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGln	402
DB	1460	CTGCAGCTTGGACAGATGATGCAATTCGCCCTTTGGATTCTTTAACGGGCAAGATGTCAG	1519
QY	403	MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal	422
DB	1520	ATGTGTCTCCGGTCTGACACTTTGTTGAAAGTGTAGGCGGATTAGCTTAATCTCTGGGTG	1579
QY	423	ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer	442
DB	1580	CGCGTGTCTTCATTTGGGAAGCTGTAGCAACCTGTGGAGGCGCATATATCATCTAGT	1639
QY	443	AspIlePheLysPheLeuLeuSer	450
DB	1640	GATATTTCAAGTTCTTGTCTGAGC	1663

RESULT 3

AAA51353
ID AAA51353 standard; cDNA; 1632 BP.
XX
AC AAA51353;
XX

DT 26-SEP-2003 (first entry)
 XX Wheat sucrose non-fermenting 4 protein cDNA from clone wll.pk0002.b3.
 DE Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.
 XX Triticum aestivum.
 OS
 XX Key Location/Qualifiers
 FH 10..1399
 FT CDS /*tag= a
 FT /product= "Sucrose_non-fermenting_4_protein"
 XX
 XX WO200036116-A2.
 XX 22-JUN-2000.
 XX 15-DEC-1999; 99WO-US029825.
 XX 16-DEC-1998; 98JS-0112564P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Allen SM, Heppard BP, Miao G, Wang Z;
 XX WPI; 2000-431594/37.
 XX P-PSDB; AAY96789.
 XX
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX
 XX Claim 18; Page 43; 48pp; English.
 XX
 XX This cDNA encodes wheat sucrose non-fermenting 4 (SNF4) protein which is
 CC involved in carbon catabolite repression in plants and seeds. The cDNA
 CC was isolated based on similarity to SNF4 proteins from Arabidopsis
 CC thaliana and Saccharomyces cerevisiae, from library wll prepared from
 CC wheat leaf from 7 day old seedling. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX
 XX SQ Sequence 1632 BP; 431 A; 370 C; 389 G; 442 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,69e-192 Length: 1632
 Score: 2031.00 Matches: 386
 Percent Similarity: 93.74% Conservative: 33
 Best Local Similarity: 86.35% Mismatches: 26
 Query Match: 86.72% Indels: 2
 DB: 3 Gaps: 1
 US-09-857-525C-2 (1-451) x AAAS1353 (1-1632)
 QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerIleu 23
 DB 2 CACGAGCGGATGTTCCAGTCGAAGGTTCGCCACTGTATTCAAGCTATCTGCAACTTA 61
 QY 24 SerProGlyIleHisGluTyrlsPheValAspGlyGluTrpArgHisGluArg 43
 DB 62 CCTCCAGGGATTTATCAGTACAGTTCAATGTGACGGCGAGTGAGCGAGGG 121
 QY 44 GlnProThrIleSerGlyGluPheGlyTleValAsnThrLeuTyrlsArgGluTyrl 63
 DB 122 CAACCTACTATACCTGGAGAGATGGGGTGTAAACACTTTATCTTGACAGGGAATTT 181
 QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83

Db 182 GACCACATAAATCTGTACTGAGCCCACTACACTGGAGCAGG-----ATGATGTG 235
 QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
 Db 236 GACAGTACAGTTTTTCAACGAATGGTTCGTTCGGATGGTGCCTTCAGGAGGTTCT 295
 QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrlsLeu 123
 Db 296 CCAAGATCTCAGAGCGTCTATACAGATCTCTAGGTGTCTGTTCGTGAGTATCTGAAT 355
 QY 124 LeuHisThrCysTyrlsAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 443
 Db 356 CGCATACAGGCTATGACCTACTACCAATCTCTGAAAGGTCAATGCTCTGCACATTAA 415
 QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
 Db 416 TTACCTGTGAAGCAATCTTTCATATCTCCATGAACAGGGGATTCCTGTGCTCTCTG 475
 QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
 Db 476 TGGGATTCATTACAGGGGTCAAGTTGTGGCCCTCTGAGCCCACTGGATTTTATCTTATA 535
 QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHisThr 203
 Db 536 TTGAGAGAGCTGGAAACTCATGGCTCAAACTCTGACAGAGAACACAGCTTGAACACACACT 595
 QY 204 IleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArg 223
 Db 596 ATATCTCGTGGAAAGAGGCTAAGCGCAAACTTATGGAAGAAATGATGGACACTTGA 655
 QY 224 ProGlnGlnHisLeuValHisAlaThrProTyrlsGluSerLeuArgAspIleAlaValLys 243
 Db 656 TCRAATCAGCATCTAGTCATGCCACCCCTTATGAATCCTTGAGGGGTATTGCCATGAAA 715
 QY 244 LeuLeuGlnAsnGlyIleSerThrValProValIleTyrlsSerSerSerAspGlySer 263
 Db 716 ATACTGGAACCTGGCATTTCTACAGTCCCAATCATCTATTCTCATCATCATCATCAT 775
 QY 264 PheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrl 283
 Db 776 TTTCGCGAGCTGTTCGATCTTCGATCCCTTTTCAGGAATTTTGAATGATCTGTAGATAC 835
 QY 284 PheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeu 303
 Db 836 TTCAGAACTCCACTGGTAGTTTGCAGTCTTAAACCAACCAACCAAGTATGCTCAATCCGCTG 895
 QY 304 GlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgPro 323
 Db 896 GGTACCTGGGTTCCAAAAATTTGGTGAACCAAAATGGTTCATCCATTGGCTATGTTCGGGCT 955
 QY 324 AsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIlePro 343
 Db 956 AATACATCTCTAGCTTCGCCCTTAATCTGTGTTCAGCTGAGGTAGTGTAGTTCATATCCC 1015
 QY 344 IleValAspAspAsnAspSerLeuLeuAspThrTyrlsSerArgSerAspIleThrAlaLeu 363
 Db 1016 ATTGTGATGATACGACTCGCTGATCGACACACTCCAGAGTGACATCACAGCTCTA 1075
 QY 364 AlaLysAspLysValTyrlsThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeu 383
 Db 1076 GCGAAGACAGAGGCTACACCCATATCGCCCTAGATGAGATGACCATTCATCAGGCCCTTG 1135
 QY 384 GlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMet 403
 Db 1136 CAGCTCGGGCAGACGCGAATTCACCTTTGGACCTTTCAATGGTCAAGATGCCAGATG 1195
 QY 404 CysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArg 423
 Db 1196 TGTCTCCAGCTGACCCCTTGTCTGAGGTTATGAGAGATTTGGCTAATCTCTGGGGTGGCT 1255
 QY 424 ArgValPheIleValGluAlaGlySerIysArgValGluGlyIleIleSerIleSerAsp 443

DB 1256 CCGGTGTCATCGTGAGGCTGGCAGCAGCGAGTGGAAAGCGTAATATCGCTGAGCGAC 1315

QY 444 IlePheLysPheLeuLeuSer 450
 DB 1316 ATATTCAGTGTCTGAGC 1336

RESULT 4
 ID AAA51351
 AC AAA51351; s-standard; cDNA; 2538 BP.

XX 26-SEP-2000 (first entry)
 XX Soybean sucrose non-fermenting 4 protein cDNA from clone sfl1.pk0004.b4.

XX Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.

XX Glycine max.
 XX Key Location/Qualifiers
 FH 395..1873
 CDS /*tag= a
 FT /product= "Sucrose non-fermenting_4_protein"

XX W0200036116-A2.
 XX 22-JUN-2000.
 XX 15-DEC-1999; 99WO-US029825.
 XX 16-DEC-1998; 98US-0112564P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Allen SM, Heppard EP, Miao G, Weng Z;
 XX WPI; 2000-431594/37.
 XX P-PSDB; AAY96787.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.

XX Claim 2; Page 39-40; 48pp; English.
 XX This cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)
 CC protein which is involved in carbon catabolite repression in plants and
 CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from library sfl1
 CC prepared from soybean immature flower. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development

XX Sequence 2538 BP; 670 A; 463 C; 571 G; 834 T; 0 U; 0 Other;
 XX Alignment Scores:
 Pred. No.: 3.94e-147 Length: 2538
 Score: 1581.00 Matches: 305
 Percent Similarity: 81.15% Conservative: 61
 Best Local Similarity: 67.63% Mismatches: 75
 Query Match: 67.51% Indels: 10
 DB: 3 Gaps: 4

US-09-857-525C-2 (1-451) x AAA51351 (1-2538)
 QY 3 GluHisLeuProMetSerProiledGlyCysProThrValPheGlnAlaIleCysSer 22

DB	530	GAGCTTCTACCGATGTCGCGAGGTTGTCACACTGTGTTCAAGTATTATAAC	589
QY	23	LeuSerProGlyIleHisGlyValAspGlyGluTrpArgHisAspGlu	42
DB	590	TTGCCACCCGGTTACCATCAGTACAAAGTTTTTTTGTGATGGAGATGGCCGCGCATGAA	649
QY	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuThrArgGlu	62
DB	650	CATCAACCTTATGTACCTACCGAATATGGATAGTCAACACTGCTCTTATTGGCCACTGAT	709
QY	63	TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp	82
DB	710	CCTAACTACATGCTGCTGTTTA-----CCTCCAGACGTTGCTTCTGGAAATAGCATGGAT	763
QY	83	ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
DB	764	GTGGATAATGATGCTTTCCCGAATGCGCGGTGGACGATGGTACTTTGAGTGAGGTG	823
QY	103	ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluThrLeu	122
DB	824	CTGCCAAGATATCAGATACCTGATGTACAAATATCCCGTCAGCGTATTTCTGCAATTCTA	883
QY	123	AsnLeuHisThrCysTyrAspLeuProAspSerGlyLysValIleAlaLeuAspIle	142
DB	884	TCTTACACACCCGCTTATGATTTACITCCCGAGTCAGCAGGTTGTGCTTGGATGTT	943
QY	143	AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro	162
DB	944	GATTCACAGTGAACAGGCAATTCATATATTCATGAGCAGGAGGATTTTCATGGCTCCT	1003
QY	163	LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu	182
DB	1064	CTTTGGGACTTCTGCAAGGGGCAATTTGTTGGTCTTCTAGTCTTCGATTTTATTATTA	1063
QY	183	IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis	202
DB	1064	ATTTTAAGAGAGCTGGGGAAATCATGGATCCCAATCTGACGAGAGGAGGCTTGAACACAT	1123
QY	203	ThrIleSerAlaTrpLysGluAlaLys-----ArgGlnThrAsnGlyArgAsn	218
DB	1124	ACCATATCAGCTTGGAAAGAGAAATCGTATCTAAATAGACAGCAATGGACATGGA	1183
QY	219	AspSerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArg	238
DB	1184	ACTGCATTTTCAAGA-----TGTTTTATCATGTCAGGCCCATATGATTAATCTGAAA	1234
QY	239	AspIleAlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSer	258
DB	1235	GATATTCGATGAAGATCTTGCAAAAGAGGTTTCACTGTTCTTATTCATTCATCT	1294
QY	259	SerSerAspGlySerPheProGlnLeuLeuHisIleAlaSerLeuSerGlyIleLeuLys	278
DB	1295	TCTGAGATGCTTCATTTCCACAGTTACTACATCTTGCTTCACTTCCAGGATACTTAA	1354
QY	279	CysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProVal	298
DB	1355	TGCATTTGTAGTATTTTAGGCACCTGCTCTAGTTCCCTTGCCCTGTACTTCAACTTCAATC	1414
QY	299	CysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeu	318
DB	1415	TGTGAATACCTGTGGCGACGTGGGTGCCAAAATTTGGGAATCAAAATCGCGGCTCTTA	1474
QY	319	AlaMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGly	338
DB	1475	GCAATGTTGAGACCAACCGCTTCTCTGCGTCAGCCCTAAATTTATTAGTTCAAGCCAA	1534
QY	339	ValSerSerIleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSer	358
DB	1535	GTAAGCTCATATACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1594
QY	359	AspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThr	378

Db	1595	GACATACAGCTTTGGCAAGAAAGAGAGATATACATATTAATCTTGACGAATGACT	1654
Qy	379	IleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGly	398
Db	1655	GTTTCATCAGGCTTTCAGCTTGGCCAGGACCATATAGTCCCTATGAG-CTTAGAAGT	1711
Qy	399	GlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAla	418
Db	1712	CAAAAGATGTCAGATGTGTGGTCTCTGATCTCTGCAATAAGTGATGGAACGCTTGSCA	1771
Qy	419	AsnProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIle	438
Db	1772	AATCCAGGTGTCAGCGGCTTGTTCATCTGGAAGCTGGGACGACGCTAGAGGCATT	1831
Qy	439	IleSerLeuSerAspIlePheLysPheLeuLeu 449	
Db	1832	GTTTCATGAGTGACATATTCAGTTCCTTCATT 1864	
RESULT 5			
ID	AAA51350		
XX	AAA51350	standard; cDNA; 2160 BP.	
AC	AAA51350;		
DT	26-SEP-2000	(first entry)	
DE	Soybean	sucrose non-fermenting 4 protein cDNA from clone ses4d.pk2040.bl.	
KW	Soybean;	sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;	
XX	carbon	catabolite repression; development; nitrogen partitioning; ss.	
OS	Glycine	max.	
PH	Key	Location/Qualifiers	
FT	185..1633		
FT	/*tag= a		
FT	/product= "sucrose_non-fermenting_4_protein"		
XX	WO200036116-A2.		
XX	22-JUN-2000.		
PF	15-DEC-1999;	99WO-US029825.	
PR	16-DEC-1998;	98US-0112564P.	
XX	(DUPO)	DU PONT DE NEMOURS & CO E I.	
PI	Allen SM,	Heppard EP, Miao G, Weng Z;	
XX	WPI; 2000-431594/37.		
DR	P-PSDB;	AA996786.	
XX	New	nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins	
PT	involved	in carbon catabolite repression in plants and seeds, useful for	
PT	controlling	carbon and nitrogen partitioning pathways during plant growth	
PT	and	development.	
XX	Claim 2;	Page 36-37; 48pp; English.	
PS	This	cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)	
CC	protein	which is involved in carbon catabolite repression in plants and	
CC	seeds.	The cDNA was isolated, based on similarity to SNF4 proteins from	
CC	Arabidopsis	thaliana and Saccharomyces cerevisiae, from library ses4d	
CC	prepared	from soybean embryonic suspension 4 days after subculture. The	
CC	polynucleotides	are used in plants to control carbon and nitrogen	
CC	partitioning	pathways during plant growth and development. The catabolite	
CC	repression	proteins would facilitate studies for better understanding the	
CC	mechanism	of catabolite repression in plants and could provide genetic	
CC	tools	to enhance or otherwise alter the accumulation of carbohydrates,	
CC	lipids	and proteins during plant growth and development	
XX	Sequence	2160 BP; 588 A; 392 C; 485 G; 693 T; 0 U; 2 Other;	

Alignment Scores:	1.06e-141	Length:	2160
Pred. No.:	1525.50	Matches:	303
Score:	80.18%	Conservative:	57
Percent Similarity:	67.48%	Mismatches:	72
Best Local Similarity:	65.14%	Indels:	17
Query Match:	3	Gaps:	6
DB:			
US-09-857-525C-2 (1-451) x AAA51350 (1-2160)			
Qy	3	GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer	22
Db	317	GAACTTCTACAAATGTCGCTGTGAAGCTGCCCAACTGTGTTTCAAGTTATTATATAGC	376
Qy	23	LeuSerProGlyIleHisGluThrLysPhePheValAspGlyGluTrpArgHisAspGlu	42
Db	377	TTGGTACCTGGTCATCATCAGTACCAAGTTTTTGTGTGAGAGATGGCGCATGATGAC	436
Qy	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAlaAsnThrLeuTyrLeuThrArgGlu	62
Db	437	CTTCAACCTTGTGAATCTGGAGATATGGAATGTTTAACTGTTTCACTGGCTACTGAT	496
Qy	63	TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp	82
Db	497	CCTAATATTTTACCTGTTTAACTCCAGACATAGTTCCTGGATCT-----AACATGGAT	550
Qy	83	ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
Db	551	GTCGACACAGCGCTTTCGACGATGTCGATTCGATGATGATGATGATGATGATGATG	610
Qy	103	ThrLeu---ArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyr	121
Db	611	TTATTGCCAAGATATCTGATGTTGATATACAGACCTCTCGTCAGCGTATTTCTGCATTC	670
Qy	122	LeuAsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAsp	141
Db	671	CTATCTATGAGTACGACATATGAATTACTTCTCGTGCAGGCAAGGTTGTACTTGGAT	730
Qy	142	IleAsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAla	161
Db	731	GTTGATCTGCTGTGAACAGGCAATTCATATCTGATGATGATGATGATGATGATG	790
Qy	162	ProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIle	181
Db	791	CCTCTCTGGGACATCTGCAAGGCGCAGTTTGTGGAGTTCTTAGTGCCTTGGATTTATT	850
Qy	182	LeuIleLeuArgGluLeuGlnThrHisGlySerAsnLeuThrGluGluGlnLeuGluThr	201
Db	851	TTAATTTTAAAGAGAGCTCGGAAATCATGGGTCCATCTAACAGAGAGAGCTTGAACA	910
Qy	202	HisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln	221
Db	911	CATACCATATCAGCTTGGAAAGGAGGAAA-----	940
Qy	222	TrpArgPro---GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIle	240
Db	941	TGGACAGGATTTACAAATGTTTATCCGTGCGAGGCGCATACAGTAATTTGAAGAGATT	1000
Qy	241	AlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSer	260
Db	1001	GCTGTGAAGATCTTGCACCAATGGAATTTCAACAGTTCCTATTATTCAT-----TCAGAA	1054
Qy	261	AspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIle	280
Db	1055	GATGTTTCAATTCACAGCTACTACATCTTGTCTTCTTCTTCTTCTTCTTCTTCTT	1114
Qy	281	CysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSer	300
Db	1115	TGCAGGATTTTAGGAATGTTCTAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTG	1174
Qy	301	IleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMet	320

1175 ATCCCTGTGGCACGTGGTCCCAAAATGGGAGTCAAAATGCCGGCTCTAGCAATG 1234
 321 LeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSer 340
 1235 TTGAGACCAATGCTTCACTTACTTTCAGCCCTAAACTTATTAGTCAAGCTCAAGTAAGT 1294
 341 SerTleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIle 360
 1295 TCAATCAATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1354
 361 ThrAlaLeuAlaLysAspLysValTyrHisValArgLeuAspGluMetThrIleHis 380
 1355 ACAGCTTGGCAAGGACAGAACTTATACCAATATTAATCTTGATGAATGAGCTGTTCAT 1414
 381 GlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArg 400
 1415 CAGCATTTGCATTTGGCCAGGATCTTATTAATCTTATGAG---CTGAGTTGTCAAGA 1471
 401 CysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnPro 420
 1472 TGTGAGATGTTTGGACACGATCTCTGCATTAAGTGAACGTTTGGCAAGTCCA 1531
 421 GlyValArgArgValPheIleValGlnAlaGlySerLysArgValGluGlyIleIleSer 440
 1532 GGTGTGAGCGGCTGTGAATTTGGAGCTGGCAGCAAGCGGTAGAAGGCATCATGCA 1591
 441 LeuSerAspIlePheLysPheLeuLeu 449
 1592 CTGAGTGACATATCAACTTCTCCIT 1618

RESULT 6
 AAA51354
 ID AAA51354 standard; cDNA; 538 BP.
 XX
 AC AAA51354;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE Maize sucrose non-fermenting 4 protein cDNA from clone csk1c.pk001.c15.
 XX
 KW Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT CDS 347..523
 FT /*tag= a
 FT /product= "partial_SNF4_protein"
 FT unsure 494..496
 FT /*tag= b
 FT /note= "Encodes Xaa which is not defined"
 XX
 PN WO200036116-A2.
 XX
 PD 22-JUN-2000.
 XX
 PF 15-DEC-1993; 99WO-US029825.
 XX
 PR 16-DEC-1993; 98US-0112564P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO B I.
 XX
 PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX
 DR WPI; 2000-431594/37.
 DR P-PSDB; AAY96789.
 XX
 PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.

PS Claim 18; Page 45; 48pp; English.
 XX This partial cDNA encodes maize (Zea mays) sucrose non-fermenting 4
 CC (SNF4) protein which is involved in carbon catabolite repression in
 CC plants and seeds. The cDNA was isolated, based on similarity to SNF4
 CC proteins from Arabidopsis thaliana and Saccharomyces cerevisiae, from
 CC library csk1c prepared from corn unpollinated developing silk 24 hours
 CC after emergence. The polynucleotides are used in plants to control carbon
 CC and nitrogen partitioning pathways during plant growth and development.
 CC The catabolite repression proteins would facilitate studies for better
 CC understanding the mechanism of catabolite repression in plants and could
 CC provide genetic tools to enhance or otherwise alter the accumulation of
 CC carbohydrates, lipids and proteins during plant growth and development
 XX
 SQ Sequence 538 BP; 146 A; 115 C; 120 G; 156 T; 0 U; 1 Other;
 Alignment Scores: Length: 538
 Pred. No.: 1,68e-82 Matches: 174
 Score: 923.00 Conservatives: 2
 Percent Similarity: 98.32% Mismatches: 3
 Best Local Similarity: 97.21% Indels: 0
 Query Match: 39.41% Gaps: 0
 DB: 3
 US-09-857-525C-2 (1-451) x AAA51354 (1-538)
 QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 DB 2 GAACATTTGGCGATGTTCTCTATCGAAGCTGCCCACTGATTTTCAGGCTATTTCGAGT 61
 QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
 DB 62 CTGTCTCTGGGATTCACGAGTACAAATTCCTTTGGAGCGGGAATGGCGGATGATGAG 121
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 DB 122 CTTCAACCTACCATATCTGGGGATTTGGCATAGTTAACACACTTTACTTGACAGGAA 181
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 DB 182 TATAACCAATAAATCAACCTTATCAAGTCCAGCACACCTGGAGGAGGATGAACATGGAT 241
 QY 83 ValAspAsnGlnAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 DB 242 GTGGATAATGAATAATTTTCAACGTACGGTTACGTTGTCAGATGCCACCGTTTCAGAGGT 301
 QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 DB 302 ACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTTCTGATATCTG 361
 QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
 DB 362 AATTTGCATACATGCTATGATTTACTCCAGATTTCTGGCAAGGTTATTGCCCTAGACAT 421
 QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 DB 422 AATTTACCTGTGAAGCAATCAATTCATATTCCTCATGAACACAGGGATTTCTGTAGTCTCT 481
 QY 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIle 181
 DB 482 CTCTGGGACTCANTCAAGGTCAATTTGGTGGGCCCCCTTAGCCCAATGGATTTCATA 538
 RESULT 7
 AAA51357
 ID AAA51357 standard; cDNA; 514 BP.
 XX
 AC AAA51357;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE Wheat sucrose non-fermenting 4 protein cDNA from clone wreln.pk0143.e2.
 XX
 KW Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

CC hybridisation or amplification technologies, as query sequences to
 CC determine homology of known sequences, as probe for use in Southern or
 CC Northern hybridisation, and to identify the presence of and/or to
 CC determine the degree of similarity between two (or more) nucleic acid
 CC sequences

XX Sequence 282 BP; 86 A; 50 C; 71 G; 75 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9.7e-37 Length: 282
 Score: 460.00 Matches: 93
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 1
 Query Match: 19.64% Indels: 1
 DB: 6 Gaps: 0

US-09-857-525C-2 (1-451) x ABL70761 (1-282)

QY 29 GlnTyrLysPheValAspGlyGluTrpArgHisAspGluArgGlnProThrIleSer 48
 DB 2 GAGTACAAATCTTTGTGGACGGGATGGCGCATGATGACGCGTCAACCTHACCATNCT 61
 QY 49 GlyGluPheGlyIleValAsnThrLeuTyrLysThrArgGluTyrAsnGlnIleAsnThr 68
 DB 62 GGGGGGTTGGCATAGTTTACACACTTTTACTTGACAGGGGAATATAACCAATAAACACC 121
 QY 69 LeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPhe 88
 DB 122 TTATCAAGTCCAAAGCACACCTTGAAGCAGGATGAACATGGATGGGATAATCAAAATTTT 181
 QY 89 GlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGlu 108
 DB 182 CAACGGGG-GTTACGTTGTGATGATGATCCGTTTCAGAAAGGTACTCTGAGATTTTCAGN 240

QY 109 AlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122

DB 241 GCTGCAATACAAATATCTAGGTGTCGTGTTCTGATATCTG 282

RESULT 9

AAA51349

ID AAA51349 standard; cDNA; 702 BP.

XX AAA51349;

AC AAA51349;

XX 26-SEP-2003 (first entry)

XX Rice sucrose non-fermenting 4 protein cDNA from clone rls6.pk0037.g9.

XX Rice; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.

XX Oryza sativa.

XX Key Location/Qualifiers

FT 3. -512

FT /*tag= a

FT /label= SNF4

FT /product= "Sucrose non-fermenting_4 protein"

FT /partial

XX WO200036116-A2.

XX 22-JUN-2000.

XX 15-DEC-1999; 99WO-US029825.

XX 16-DEC-1998; 98US-0112564P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Allen SM, Heppard EP, Miao G, Weng Z;

XX WPI; 2000-431594/37.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.

XX Claim 2; Page 35-36; 40pp; English.

XX This cDNA encodes rice (*Oryza sativa*) sucrose non-fermenting 4 (SNF4)
 CC protein which is involved in carbon catabolite repression in plants and
 CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
 CC rls6.pk0037.g9 prepared from rice leaf 15 days after germination, 6 hours
 CC after infection of strain Magaporthe grisea 4360-R-67 (AVR2-YAMO)
 CC (susceptible). The polynucleotides are used in plants to control carbon
 CC and nitrogen partitioning pathways during plant growth and development.
 CC The catabolite repression proteins would facilitate studies for better
 CC understanding the mechanism of catabolite repression in plants and could
 CC provide genetic tools to enhance or otherwise alter the accumulation of
 CC carbohydrates, lipids and proteins during plant growth and development

XX Sequence 702 BP; 233 A; 136 C; 155 G; 178 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.27e-29 Length: 702
 Score: 392.00 Matches: 84
 Percent Similarity: 62.11% Conservative: 34
 Best Local Similarity: 44.21% Mismatches: 64
 Query Match: 16.74% Indels: 8
 DB: 3 Gaps: 2

US-09-857-525C-2 (1-451) x AAA51349 (1-702)

QY 261 AspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIle 280

DB 18 GATTCATCAGGATGCGTTTGTCTTGCTTGACCCCTCCAGGATTTGTAATTATT 77

QY 281 CysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSer 300

DB 78 TGCTCAAGCTCGAAGACAGCGCTGAAGGGTACTCATTTCTGCAACACAGATTGTCACT 137

QY 301 IleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMet 320

DB 138 ATGCTTATTTGGTACATGTTGTCACCATACTGCGAAGCAAGCAATAGACAGCTTAGA 197

QY 321 LeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSer 340

DB 198 TCGCGACCAAGCACTCTCTTAATTCATGCTGGATTGCTGCTGAAGATAGTAGTAGC 257

QY 341 SerIleProIleValAspAspAsnSerLeuLeuAspThrTyrSerArgSerAspIle 360

DB 258 TCAATTCCTATAGTTGACGATAATGCGCTCTCTTGATGCTACTCGCTCAGTGATATC 317

QY 361 ThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHis 380

DB 318 ATGGCTCTAGGCAAGAATGATGCTACCTCGTATTGAGCTTCAACAGGCGAGGTGAG 377

QY 381 GlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArg 400

DB 378 CATGCCCTTGAGCTGCAATACACAGGTG-----ANTGCCGAGGA 416

QY 401 ---CysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsn 419

DB 417 CACTGTCTACCTGCTTGAGCACTAGTACCTTCTGGAGGTTTGGAGCAATTGTGAGCT 476

QY 420 ProGlyValArgValPheIleValGlnAlaGlySerLysArgValGluGlyIleIle 439

DB 477 CCAGGGGTGGCGGAGTCTGCTTATTGAACCAAGGAGCAGATTGTGCAAGGAATATC 536

QY 440 SerLeuSerAspIlePheLysPheLeuLeu 449

DB 537 TCATTGAGGACGCAATTACATTTCTCAT 566

RESULT 10
 ID AAL48614
 XX AAL48614 standard; cDNA; 1167 BP.
 AC AAL48614;
 XX
 DT 11-OCT-2002 (first entry)
 XX
 DE Human insulin receptor signaling modifier cDNA SEQ ID NO: 11.
 DE
 DE Human insulin receptor signaling; insulin receptor signaling modifier;
 KW Human; insulin receptor signaling; insulin receptor signaling modifier;
 KW ISM; diabetes; metabolic syndrome; antidiabetic; gene; ss.
 KW
 XX
 OS Homo sapiens.
 OS
 PN WC200255664-A2.
 PN
 PD 18-JUL-2002.
 PD
 XX
 PF 11-JAN-2002; 2002WO-US001048.
 PF
 XX 12-JAN-2001; 2001US-0261226P.
 PR 12-JAN-2001; 2001US-0261303P.
 PR 12-JAN-2001; 2001US-0261304P.
 PR 12-JAN-2001; 2001US-0261335P.
 PR 12-JAN-2001; 2001US-0261336P.
 PR 12-JAN-2001; 2001US-0261361P.
 PR 12-JAN-2001; 2001US-0261456P.
 PR 12-JAN-2001; 2001US-0261457P.
 PR 12-JAN-2001; 2001US-0261458P.
 PR 12-JAN-2001; 2001US-0261459P.
 PR 12-JAN-2001; 2001US-0261461P.
 PR 12-JAN-2001; 2001US-0261518P.
 PR 12-JAN-2001; 2001US-0261531P.
 PR 12-JAN-2001; 2001US-0261532P.
 PR 12-JAN-2001; 2001US-0261589P.
 PR 12-JAN-2001; 2001US-0261590P.
 PR 12-JAN-2001; 2001US-0261694P.
 PR 12-JAN-2001; 2001US-0261695P.
 PR 12-JAN-2001; 2001US-0261697P.
 XX
 PA (EXEL-) EXELIXIS INC.
 PA
 PI Seidel-Dugan C, Ferguson KC, Kidd T;
 PI WPI; 2002-399664/64.
 DR P-PSDB; AA018495.
 DR
 XX Identifying an insulin receptor signaling modulator, useful as drug
 PT targets for treating diabetes or metabolic disorders, comprises
 PT contacting an assay system comprising insulin receptor signaling
 PT modifiers with a test agent.
 PT
 XX Disclosure; Page 48; 232pp; English.
 PS
 XX The present invention relates to a method of identifying a candidate
 CC insulin receptor (INR) signaling modulating agent, involving contacting
 CC an assay system comprising an insulin receptor signaling modifier (ISM)
 CC polypeptide or nucleic acid with a test agent, and detecting a test agent
 CC biased activity of the assay system. The method is useful for
 CC identifying candidate INR signaling modulating agents. ISM genes may be
 CC used as drug targets for treatment of disorders related to INR signaling
 CC such as diabetes or metabolic syndrome. ISM nucleic acids and
 CC polypeptides are useful for identifying and testing agents that modulate
 CC ISM function and for other applications related to the involvement of ISM
 CC in INR signaling, and for identifying subjects having a predisposition to
 CC such diseases associated with INR signaling. The present sequence is an
 CC ISM coding sequence described in the exemplification of the invention
 CC
 SQ Sequence 1167 BP; 352 A; 246 C; 265 G; 304 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 D-nd Wc. 3 176-27 Length: 11167

```

D5 858 G-----TATTTGAGGT-----GTTGTGAGTGCATTAAGCTGGA 893
QY 408 pProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgValPheIleVa 428
D5 894 AATACTGGAGACCATCGTGACAGATAGTAAGAGCTGAGGTCCATCGGCTGGTGGT 953
QY 428 1---GluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPh 447
D5 954 AATATGAGCAGATAGT-----ATTGTGGTATATTATTCCTCGGACATTCGCAAGC 1007
QY 447 eLeuLeu 449
D5 1008 CCTGATC 1014

RESULT 11
AA006882
ID AA006882 standard; cDNA; 1435 BP.
XX
AC AA006882;
XX
D5 26-APR-1999 (first entry)
XX
D5 Disease associated protein kinase DAPK-7 cDNA.
XX
DAPK-7; disease associated protein kinase; human; diagnosis; therapy;
XX adult respiratory distress syndrome; allergy; asthma; arteriosclerosis;
XX bronchitis; emphysema; hyperesinophilia; myocardial inflammation;
XX pericardial inflammation; anaemia; rheumatoid arthritis;
XX Addison's disease; AIDS; atherosclerosis; atopic dermatitis;
XX dermatomyositis; diabetes mellitus; glomerulonephritis; gout;
XX Grave's disease; lupus erythematosus; multiple sclerosis;
XX myasthenia gravis; osteoarthritis; osteoporosis; pancreatitis;
XX polycystic kidney disease; polymyositis; scleroderma;
XX Sjogren's syndrome; autoimmune thyroiditis; cancer; infection; trauma;
XX cell proliferation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 265..1503
FT FT /*tag= a
FT
PN WC9858052-A2.
XX
XX 23-DEC-1998.
XX
XX 18-JUN-1998; 98WO-US012813.
XX
XX 19-JUN-1997; 97US-00873989.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Barčman O, Hillman JL, Corley NC, Guegler KJ, Lal P, Goli SK,
XX Shah P;
XX
XX WPI; 1999-080952/07.
XX
XX P-PSDB; AAW88438.
XX
XX New disease associated protein kinases - used to stimulate cell
XX proliferation and to treat the immune response and cancer.
XX
XX Claim 5; Page 66-67; 93pp; English.
XX
XX This cDNA sequence codes for human disease associated protein kinase DAPK
XX -7 (see AAW88438). DAPK-7 cDNA was first identified in the PENITUT01 cDNA
XX library using a computer search for amino acid alignments, and a
XX consensus sequence was derived from the extended and overlapping Incyte
XX clones 3075712/HEARNOT01, 842220/PROSTUT05, 1364747/SCORN002, 145972 and
XX 145802/PENITUT01 and 1479332/CORPNOT02. DAPK-7 shows 73% homology with
XX the human foetal liver AMPK gamma subunit (GI 335856), and is associated
XX with cDNA libraries which are immortalised or cancerous and show
XX inflammatory or immune responses. The invention provides disease

```

```

CC associated protein kinases DAPK-1 to DAPK-7 (see AAW88432-38) and cDNA
CC clones encoding them (see AAX06831-36 and AAX06882), as well as
CC expression vectors, host cells, agonists, antagonists and antibodies. The
CC invention further provides uses of such products in the diagnosis,
CC prevention and treatment of diseases associated with cell proliferation,
CC especially cancer or an immune response (claimed). Conditions that may be
CC treated include adult respiratory distress syndrome, allergies, asthma,
CC arteriosclerosis, bronchitis, emphysema, hyperesinophilia, myocardial or
CC pericardial inflammation, rheumatoid arthritis, Addison's disease, AIDS,
CC anaemia, atherosclerosis, various diseases of the digestive system,
CC atopic dermatitis, dermatomyositis, diabetes mellitus,
CC glomerulonephritis, gout, Grave's disease, lupus erythematosus, multiple
CC sclerosis, myasthenia gravis, osteoarthritis, osteoporosis, pancreatitis,
CC polycystic kidney disease, polymyositis, scleroderma, Sjogren's
CC syndrome, autoimmune thyroiditis, complications of cancer, extracorporeal
CC circulation, viral, bacterial, fungal, parasitic, protozoal and
CC helminthic infections, and trauma (disclosed). The DAPK nucleic acids are
CC also used in a method for detection of DAPK expression levels in a
CC biological sample
XX
SQ Sequence 1435 BP; 421 A; 298 C; 331 G; 385 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.: 5.13e-27 Length: 1435
Score: 373.00 Matches: 104
Percent Similarity: 51.29% Conservative: 75
Best Local Similarity: 29.80% Mismatches: 128
Query Match: 15.93% Indels: 42
DB: 2 Gaps: 13

```

US-09-857-525C-2 (1-451) x AAX06882 (1-1435)

```

QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgVal---SerGluTyrLeu 122
D5 140 CTGGAGTTCGAGGACGAGCAGTACAGTCTCCACACGTTCTTACATCGATTCATG 199
QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
D5 200 AGGTCAACAAGTGTATGATCATCGTTCACACGTTCTTCAAGCTGTGTCTTCTACT 259
QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
D5 260 ACATTACAAGTTAAAGAGCCCTCTTCTTGTAGCCACCGGTCCGAGCAGCGCCA 319
QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
D5 320 CTGTGGGAGAGTAAAGAACAAAGTTTGTAGGAATGCTAAACATTACAGATTTCATAAT 379
QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGlu---GluGlnLeuGluThr 201
D5 380 ATACTACATAGATAC-----TATAAATCACCTATGTTACAGATTATGAATTAGAGAA 433
QY 202 HisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln 221
D5 434 CATAAATTCGAACATCGAGGAGCCTTTATTATCAA-----GAAACA 475
QY 222 TrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAla 241
D5 476 TTTAAGCCT-----TTAGTGAATATATCTCAGATGCAAGCCTTTCGATCTCTA 526
QY 242 ValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAsp 261
D5 527 TACTCTTGTATCAAAAATAAATCCACAGATTGCCCGTTATTGACCTTATCAGTGGGAAT 586
QY 262 GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
D5 587 GCA-----CTTTATATCTACCCCAAAAGAACTCTCAAGTTCCTC--- 628
QY 282 ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
D5 629 CAGCTTTTATGTCGTATATGCCAAAGCTTCCTTCATGAAGCAGAACTGATGAGCTT 688
QY 302 ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321

```

```

Db 689 GGAATAGGAGCTAC-----CACAACATGGCTTCATA 721
Qy 322 ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
Db 722 CATCCAGACATCCCATCATCAAGCCTTGAACATATTGTGGAAAGACGAATATCAGCT 781
Qy 342 IleProIleValAspAsnAspSerLeuLeuAspThrThrLeuSerSerAspIleThr 361
Db 782 CTGCCTGTGTGGATGAGTCAGGAAAGTTGTAGATATTATTCACAAATTTGATGTAATT 841
Qy 362 AlaLeuAlaLeuAspLysValThrHisValArgLeuAspGluMetThrIleHisGln 381
Db 842 AATCTTGCTGCTGGAACAAACATCAATACCTA-----GATATCAGGTGACCCAG 892
Qy 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
Db 893 GCCCTTCAGCACCGCTTCACAG-----TATTTGAAGGT-----GTT 928
Qy 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly 421
Db 929 GTGAAGTCGATAAGCTGGAATACTGGAGACCCTGGAGACGATGAGATGATAGAGCTGAG 988
Qy 422 ValArgArgValPheIleVal---GluAlaGlySerLysArgValGluGlyIleIleSer 440
Db 989 GTCCATCGCTGGTGGTGGTAAATGAGCAGATAGT-----ATTGTGGGTATTATTCC 1042
Qy 441 LeuSerAspIlePheLysPheLeuLeu 449
Db 1043 CTGTCGACATCTCGAAGCCCTGATC 1069

RESULT 12
AAH14839
ID AAH14839 standard; cDNA; 2223 BP.
AC AAH14839;
DT 26-JUN-2001 (first entry)
DX Human cDNA sequence SEQ ID NO:12660.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000JP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX Claim 8; SEQ ID NO 12660; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:

```

```

CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence and sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2223 BP; 633 A; 431 C; 496 G; 663 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.:          Length:          2223
Score:              373.00           Matches:      104
Percent Similarity: 51.20%           Conservative: 75
Best Local Similarity: 29.80%         Mismatches:   128
Query Match:        15.93%           Indels:       42
DB:                 4                Gaps:        13

```

US-09-857-525C-2 (1-451) x AAH14839 (1-2223)

```

Qy 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgVal---SerGluTyrLeu 122
Db 172 CTGGAGTTCGAGGAGCAAGCAGTAGAAGACTCAGAAAGTGGTGTTCATGCGATTCATG 231
Qy 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 232 AGCTCACACACAGCTTATGACATCGTTCACACAGTTCACAGCTTGTGCTTTGTACT 291
Qy 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db 292 ACATTACAAGTTAAAAAGGCTTCTTTGCTTTGGTAGCCACCGTTCGACGAGCGCCCA 351
Qy 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 352 CTGTGGGAGAGTAAAAACAAAGTTTGTAGGAATGCTACACATTACACATTTCATAAT 411
Qy 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGlu---GluGlnLeuGluThr 201
Db 412 ATACTACATAGATAC-----TATAAATCACCTATGTTACAGATTATGAATTAGAGGA 465
Qy 202 HisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln 221
Db 466 CATAAATTTGAACATCGAGGGAGCTTTATTACAA-----GAAACA 507
Qy 222 TrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAla 241
Db 508 TTTAAGCCT-----TTAGTGAATATATCTCCAGATGCAAGCTCTTCGATGCTGTA 558
Qy 242 ValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAsp 261
Db 559 TACTCTCTTATCAAAAATAAATCCACAGATGTCGCGCTTATTGACCTATCATGTGGGAA 618
Qy 262 GlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
Db 619 GCA-----CTTTATATACATTACCAAAAGATCTCCTCAAGTTCCTC--- 660
Qy 282 ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
Db 661 CAGCTTTTATGCTCATATGCCAAGCTGCTTCATGACGAGAACCTGGATGAGCTT 720

```



```
Db 619 GCA-----CTTATATACCTTACCCCAAAAGATCCTCAAGTCTC--- 660
Qy 282 ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
Db 661 CAGCTTTTATGCTGATATGCAAGGCTGCTTTCATGAAGCAGAACTGATGAGCTT 720
Qy 302 ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321
Db 721 GGAATAGGAAGCTAC-----CACAACTTGCCTTCATA 753
Qy 322 ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
Db 754 CATCCAGACACTCCCATCATCAAGCCCTTGAACATATTGTGGAAAGACGAAATACCT 813
Qy 342 IleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr 361
Db 814 CTGCTGTGTGGATGAGTCAGAAAGTTGTAGATATTATTCCAAATTTGATGTAAAT 873
Qy 362 AlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGln 381
Db 874 AATCTTGCTGAGAAACATACATAACCTA-----GATATACGGTGACCCAG 924
Qy 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
Db 925 GCCCTTCAGCCGTTCCAG-----TATTTGAAGT-----GTT 960
Qy 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly 421
Db 961 GTGAAGTGCAATAAGCTGGAATACTGGAGACCATCGTGGACAGAATAGTAGAGCTGAG 1020
Qy 422 ValArgGlnValPheIleVal---GluAlaGlySerLysArgValGluGlyIleIleSer 440
Db 1021 GTCATATCGCTGGTGGTGGTAAATGAGCAGATAGT-----ATTGGGTATTATTCC 1074
Qy 441 LeuSerAspIlePheLysPheLeuLeu 449
Db 1075 CTGTCGGACATCTGCAAGCCCTGATC 1101
RESULT 14
AAL48616
ID AAL48616 standard; cDNA; 1578 BP.
AC AAL48616;
XX
DT 11-OCT-2002 (first entry)
DE Human insulin receptor signaling modifier cDNA SEQ ID NO: 15.
XX
KW Human; insulin receptor signaling; insulin receptor signaling modifier;
XX
OS Homo sapiens.
XX
PN WC200255664-A2.
XX
PD 18-JUL-2002.
XX
PF 11-JAN-2002; 2002WO-US001048.
XX
PR 12-JAN-2001; 2001US-0261226P.
PR 12-JAN-2001; 2001US-0261303P.
PR 12-JAN-2001; 2001US-0261304P.
PR 12-JAN-2001; 2001US-0261335P.
PR 12-JAN-2001; 2001US-0261336P.
PR 12-JAN-2001; 2001US-0261361P.
PR 12-JAN-2001; 2001US-0261456P.
PR 12-JAN-2001; 2001US-0261457P.
PR 12-JAN-2001; 2001US-0261458P.
PR 12-JAN-2001; 2001US-0261459P.
PR 12-JAN-2001; 2001US-0261461P.
PR 12-JAN-2001; 2001US-0261518P.
PR 12-JAN-2001; 2001US-0261531P.
```

```
PR 12-JAN-2001; 2001US-0261532P.
PR 12-JAN-2001; 2001US-0261589P.
PR 12-JAN-2001; 2001US-0261590P.
PR 12-JAN-2001; 2001US-0261694P.
PR 12-JAN-2001; 2001US-0261695P.
PR 12-JAN-2001; 2001US-0261697P.
XX (EXEL-) EXELIXIS INC.
XX
PI Seidel-Dugan C, Ferguson KC, Kidd T;
XX
DR WPI: 2002-599564/64.
XX P-PSDB; AAO18497.
PT Identifying an insulin receptor signaling modulator, useful as drug
PT targets for treating diabetes or metabolic disorders, comprises
PT contacting an assay system comprising insulin receptor signaling
PT modifiers with a test agent.
XX
PS Disclosure; Page 53-54; 232pp; English.
XX
CC The present invention relates to a method of identifying a candidate
CC insulin receptor (INR) signaling modulating agent, involving contacting
CC an assay system comprising an insulin receptor signaling modifier (ISM)
CC polypeptide or nucleic acid with a test agent, and detecting a test agent
CC biased activity of the assay system. The method is useful for
CC identifying candidate INR signaling modulating agents. ISM genes may be
CC used as drug targets for treatment of disorders related to INR signaling
CC such as diabetes or metabolic syndrome. ISM nucleic acids and
CC polypeptides are useful for identifying and testing agents that modulate
CC ISM function and for other applications related to the involvement of ISM
CC in INR signaling, and for identifying subjects having a predisposition to
CC such diseases associated with INR signaling. The present sequence is an
CC ISM coding sequence described in the exemplification of the invention
XX
SQ Sequence 1578 BP; 400 A; 377 C; 379 G; 422 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,348-26 Length: 1578
Score: 367.00 Matches: 111
Percent Similarity: 48.43% Conservative: 74
Best Local Similarity: 29.06% Mismatches: 133
Query Match: 15.67% Indels: 64
DB: 6 Gaps: 14
US-09-857-525C-2 (1-451) x AAL48616 (1-1578)
Qy 69 LeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPhe 88
Db 107 ATTTCTTCAGATAGTCCCCAGCT-----GTGAAATAGCATCTCT 148
Qy 89 GlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGlu 108
Db 149 CAAGAGACCCAGAA---TCCAACAATAGCGTG----- 178
Qy 109 AlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyr 128
Db 179 -----TATACTCTTCATGAGTCTCATCGTCTAT 211
Qy 129 AspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProValLysGln 148
Db 212 GACCTGATTCACCAAGCTCCAAATTTGGTTGATTGATACGCTCCCTGAGGTGAAGAA 271
Qy 149 SerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTyrAspSerPheArg 168
Db 272 GCTTTTTCCTTTCCTGACTAACCGTGTACGAGTCCCTTCCTTATGGGATAGTAGAAG 331
Qy 169 GlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGlu 188
Db 332 CAAAGTTTTTGGGCGATGCTGACCATCACTGATTTTCATCAATATCTTCGACCGCTAC--- 388
Qy 189 ThrHisGlySerAsnLeuThrGlu---GluGlnLeuGluThrHisThrIleSerAlaTrp 207
```

Db 389 ---TATTAATCAGCCTTGGTACAGATCTATGAGCTAGAGACACACAGATGAACTTGG 445
 QY 208 LysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHis 227
 Db 446 AGAGAGGTGTATCTCCAG-----GACTCTTTTAAACCG----- 478
 QY 228 LeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuGlnAsn 247
 Db 479 CITGTCTGCATTTCTCTTAATGCCAGCTTGTTGATGCTGCTCTTCTCATTAATTCGGAAC 538
 QY 248 GlyIleSerThrValProValIleTyrSerSerSerSerSerSerSerSerSerSer 267
 Db 539 AAGATCCACAGGCTGCCAGTTATGACCCAGATCAGCAAT-----ACT 583
 QY 268 LeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSer 287
 Db 584 TTGTATCATCTCCACCAAGCCCATCTGAAGTTCTCTC---AAATTGTTATCATCTGAG 640
 QY 288 ThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpVal 307
 Db 641 TTCCCAAGCCAGATTCATGTCTCAAGTCTCTGGAAGAGCTACAGATGCACTATGCC 700
 QY 308 ProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeu 327
 Db 701 -----AATATTGCTATGTTGCTGCACTACCAACCCCGTC 733
 QY 328 SerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAsp 347
 Db 734 TATGTGGCTCTGGGGATTTTGTACAGATCTGAGTCTCAGCCCTGCCAGTGGTGATGAG 793
 QY 348 AsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaLysAspLys 367
 Db 794 AAGGGCGTGTGGTGCATCTACTCAAGCTTGAATGTTATCATCTGGCAGCAGAAAG 853
 QY 368 ValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGln 387
 Db 854 ACCTACAACAACCTA-----GATGTATCTGTGACTAAGCCCTTGCACATCGATCA 904
 QY 388 AspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSer 407
 Db 905 CAT-----TACTTTGAGGT-----GTTCTCAAGTCTCTACCTGCAT 940
 QY 408 AspProLeuLysValMetGluArgLeuAlaAsnProGlyValArgValPheIle 427
 Db 941 GAGACTCTGGAGACCATCATCAAGGCTAGTGGAGCAGAGGTTACCGACTGTGATG 1000
 QY 428 ValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPhe 447
 Db 1001 GTGATGAAATGATGTG---GTCAAGGGAATGTATCACTGTCTGACATCTCTGCAGGCC 1057
 QY 448 LeuLeu 449
 Db 1058 CTGGTG 1063
 RESULT 15
 ID ABK84324
 XX AC ABK84324;
 XX DT 14-AUG-2002 (first entry)
 DE Human cDNA differentially expressed in granulocytic cells #895.
 XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.

OS Homo sapiens.
 PN WO200228999-A2.
 XX 11-APR-2002.
 XX 03-OCT-2001; 2001WO-US030821.
 PF 03-OCT-2000; 2000US-0237189P.
 PR (GENE-) GENE LOGIC INC.
 PA Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 PI WPI; 2002-435328/46.
 DR Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.
 XX Claim 1; SEQ ID NO 895; 114pp; English.
 XX The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) GA by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject, or
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease, also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pot_sequences
 XX SQ Sequence 1578 BP; 400 A; 377 C; 379 G; 422 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,34e-26 Length: 1578
 Score: 367.00 Matches: 111
 Percent Similarity: 48.43% Conservative: 74
 Best Local Similarity: 29.06% Mismatches: 133
 Query Match: 15.67% Indels: 64
 DB: 6 Gaps: 14

US-09-857-525C-2 (1-451) x ABK84324 (1-1578)

QY 69 LeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAspPhe 88
 Db 107 ATTCTTCAGATAGTCCCGCAGCT-----GTGGAAATGAGCATCT 148

Qy 89 GlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGlu 108
 Db 149 CAAGAGACCCAGAA---TCCACATAGCGT----- 178
 Qy 109 AlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyr 128
 Db 179 -----TATACCTTCCTCATGAAGTCTCATCGCTGCTAT 211
 Qy 129 AspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProValLysGln 148
 Db 212 GACCTGATCCACAGCTCCAAATGGTTGATTGATAGCTCCCTGCAGGTGAAGAA 271
 Qy 149 SerPheHisIleLeuHisGlnGlnGlyIleProValAlaProLeuTrpAspSerPheArg 168
 Db 272 GCTTTTGTGCTTTGGTGACTAACGGGTGTACAGCTGCCCTTTATGGGATAGTAAGAAG 331
 Qy 169 GlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGlu 188
 Db 332 CAAAGTTTGTGGGATGCTGACCATCATCTGATGTTTCATCAATNTCCCTGCACCGCTAC--- 388
 Qy 189 ThrHisGlySerAsnLeuThrGlu---GluGlnLeuGluThrHisThrIleSerAlaTrp 207
 Db 389 ---TATAAATCAGCCTTGGTACAGATCTATGAGCTAGAGAACACACAGATAGAAACTTGG 445
 Qy 208 LysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHis 227
 Db 446 AGAGAGGTGTATCTCCAG-----GACTCCTTTAAACOG----- 478
 Qy 228 LeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeuGlnAsn 247
 Db 479 CTGTCTGCATTTCTCCCAAAGCCAGCTTGTGATGCTGCTCTCTCATTAATTCGGAC 538
 Qy 248 GlyIleSerThrValProValIleTyrSerSerSerAspGlySerPheProGlnLeu 267
 Db 539 AAGATCCACAGGCTGCCAGTTATTGACCCAGAAATCAGGCAAT-----ACT 583
 Qy 268 LeuHisIleAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSer 287
 Db 584 TTGTACATCTCTCCCAACAGCGCATCTCTGAAGTTCTTC---AAATGTTTTATCACTGAG 640
 Qy 288 ThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpVal 307
 Db 641 TTCCCAAGCCAGAGTTCATGTCGAAGTCTCGAGAGCTACAGATTGGCACCTATGCC 700
 Qy 308 ProLysIleGlyAspLeuAsnSerArgProLeuAlaValMetLeuArgProAsnAlaSerLeu 327
 Db 701 -----AATATTGCTATGTTGCACTACCAACCCCGCTC 733
 Qy 328 SerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAsp 347
 Db 734 TATGTGGCTCTGGGATTTTGTACAGCATGTCAGTCTCAGCCCTGCGAGTGTGATGAG 793
 Qy 348 AsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaLysAspLys 367
 Db 794 AAGGGCGCTGTGGTGACATCTACTCCAAGTTGATTTATCAATCTGCGACAGCAAAAAG 853
 Qy 368 ValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGln 387
 Db 854 ACCTACAAACACCTA-----GATGTATCTGTGACTAAAGCCTTGCAACATCGATCA 904
 Qy 388 AspAlaAsnThrProPhePheAsnGlnArgCysGlnMetCysLeuArgSer 407
 Db 905 CAT-----TACTTTGAGGGT-----GTTCTCAAGTGTACTCTGAT 940
 Qy 408 AspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgValPheIle 427
 Db 941 GAGACTCTGGAGACCATCATCAACAGGTAGTGGAGAGCAGAGGTTCAACCGACTGTAGTG 1000
 Qy 428 ValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPhe 447
 Db 1001 GTGGATGAAATGATGTG---GTCAAGGGAATTGTATCACTGCTGACATCTCTGCAGGCC 1057

Qy 448 LeuLeu 449
 Db 1058 CTGGTG 1063
 Search completed: July 9, 2004, 08:38:22
 Job time : 628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CV protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2004, 07:01:49 ; Search time 3645 Seconds
(without alignments)
3694.882 Million cell updates/sec

Title: US-09-857-525C-2
Perfect score: 2342
Sequence: 1 TREHLPSPBIGCTVQAL.....SKRVEGIISLSIDIKFKLLSL 451

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09857525/runat_07072004_161018_305/app_query.fasta_1.647
-DB=EST -QMT=fastcap -SUFFI=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRP=US09857525 @CGN 1.1 3437 @runat_07072004_161018_305 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_yrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2107	90.0	2227	11	AY109512	Zea mays
2	1274	54.4	873	14	CD434895	EL01N0330
3	1090	46.5	666	14	CA2111312	SCRLAD109
4	1087	46.4	644	10	AW744961	LGI_385_D
5	1086	46.4	738	14	CD878389	WHE4018_C
6	1064	45.4	737	14	CA500333	WHE4018_C
7	1050	44.8	712	14	CD903099	WHE4018_C
8	1045	44.6	683	14	CA262684	SCPLB202
9	1018	43.5	601	13	BU098591	SCPLB202
10	999	42.7	595	10	BE598304	PII_68_G0
11	969	41.4	577	14	CA218557	SCUFADIC0
12	964	41.2	559	10	BE123279	945040E06
13	963	41.1	549	13	BU499410	946174G09
14	953	40.7	631	14	CA253676	SCRLFLA10
15	943	40.3	831	10	BE704511	SC01_08c1
16	939	40.1	558	10	BE598914	PII_84_E0
17	913	39.0	726	14	CD902160	G356_156C
18	904	38.6	594	12	BU450369	BU450369
19	898	38.3	604	12	BJ477917	BJ477917
20	892	38.1	690	13	BQ163922	952081C10
21	891	38.0	534	12	BI423527	949049B11
22	889	38.0	563	12	BG605114	WHE2327_E
23	889	38.0	680	13	CA148077	SCEBRZ101
24	875	37.4	911	13	CA067477	SCQAD106
25	867	37.0	626	14	CA498139	WHE3239_C
26	855	36.5	612	14	CA595438	wpaic.p40
27	848.5	36.2	596	14	CF052805	QCM8C01_Y
28	838	35.8	689	14	CF635900	ZMRW00_0
29	835	35.7	660	14	CB008818	VVC060F11
30	828	35.4	693	14	CD863501	AZ01.106P
31	824	35.2	571	14	CF302989	ABFI_01-
32	814.5	34.8	745	14	CF438706	EST675051
33	814	34.8	700	13	CA124616	SCQGLR108
34	810.5	34.6	707	14	CF087439	OHM14P15
35	809	34.5	549	13	BU969888	HE12P01r
36	806	34.4	607	14	CK103865	1002P24.5
37	800	34.2	541	12	BJ303942	BU303942
38	794.5	33.9	626	14	CD995954	QBB7C06_X
39	793.5	33.9	780	12	BI930509	EST550398
40	787	33.6	555	12	BI479562	WHE3455_B
41	785	33.5	530	14	CF325983	JMT11_04-
42	776.5	33.2	716	13	BU025775	QHG11F23
43	769	32.8	712	10	BF625365	HV3NEA000
44	767	32.7	473	10	AW676933	DGI_2_C07
45	767	32.7	524	12	BG560098	RHI22_63_

ALIGNMENTS

RESULT 1
AY109512
LOCUS AY109512 2227 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL367_1 mRNA sequence.
ACCESSION AY109512
VERSION AY109512.1 GI:21213260
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2227)

AUTHORS Hainey,C.F., Dolan,M., Xiao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Banafey,M., Morgante,M. and Tingey,S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo probes
 JOURNAL Unpublished (2002)
 AUTHORS 2 (bases 1 to 2227)
 TITLE Direct Submission
 JOURNAL Coe,E.H.
 COMMENT Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
 If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
 FEATURES
 source
 1..2227
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="WaizeDB:631725"
 /db_xref="Taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN
 Alignment Scores:
 Pred. No.: 2,47e-236 Length: 2227
 Score: 2107.00 Matches: 412
 Percent Similarity: 92.87% Conservative: 5
 Best Local Similarity: 91.76% Mismatches: 32
 Query Match: 89.97% Indels: 0
 DB: 11 Gaps: 0

US-09-857-525c-2 (1-451) x AY109512 (1-2227)

QY 3 GluHisLeuProMetSerProFleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 |||||
 DB 640 GAACATTTGCCAGTCTCTCTGTCGAGGCTGCCCTGATTTTCAGGCTATTTCAGGC 699
 |||||

QY 23 LeuSerProGlyIleHisGluTyrLysPheValAspGlyGluTyrArgHisAspGlu 42
 |||||
 DB 700 CTGCTCCAGGATTTCATGATGACAAATCTATGTGGACGGGAGTGGGCTCATGATGAG 759
 |||||

QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 |||||
 DB 760 CGCAACCTACTATATCTGCGGAGTTCGGTATAGTTAAACACATATCTGACAAAGGAA 819
 |||||

QY 63 TyrAsnGlnIleAsnThrLeuSerProSerThrProGlySerArgMetAsnMetAsp 82
 |||||
 DB 820 TTTAATCAAAATAATTCCTTTATTAATCCAGCACACCTGGAGCGAGTAACATGAT 879
 |||||

QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 |||||
 DB 880 GTGGATAATGAAATTTTCAACGTACGGTTACGTTGTCGATGTTACCGTTTCAGAGGT 939
 |||||

QY 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 |||||
 DB 940 ACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTGCTGTTTCTGAGTATCTG 999
 |||||

QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
 |||||
 DB 1000 AATTGCAATACGCTATGATTCTCCGAGATTCTGCCAAGGTATTGCTCCCTGACATT 1059
 |||||

QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 |||||

DB 1060 AATTACTCTGTGAAGCAATCAATTCATATTCCTCCATGAACAGGGGATTCCTGTAGTCCT 1119
 QY 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 |||||
 DB 1120 CTCCTGGAGCTCATTCAGAGGTCAATTTGTGGCTCTTCGAGCCCATTCATCTACTT 1179
 |||||

QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
 |||||
 DB 1180 ATATTTCGGGAGCTAGAACTCATGGCTCGAACTTTCAGAGAAGATCAGCTTGAACACAC 1239
 |||||

QY 203 ThrIleSerAlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyr 222
 |||||
 DB 1240 ACTATATCTGCATGGAAGAGGCTAAGCGCAAACTTGTGGAAGAAATGATGTCACTGG 1299
 |||||

QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
 |||||
 DB 1300 CGAGCACATCAGCATCTAGTGCATGCCACCTTATGAGTCTCTGAGGACATTCAGTA 1359
 |||||

QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
 |||||
 DB 1360 AAGCTTTTGTCTAAATGACATTTCTACAGTGCAGTATTTATTCATCATCAGATGGA 1419
 |||||

QY 263 SerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
 |||||
 DB 1420 TCATTCCCTCAGTATTGACCTTCATCCCTTTCTGGAATTTTNNNNNNNNNNNN 1479
 |||||

QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 |||||
 DB 1480 NNNNNNNNNNTCACTGTAATTTGCTATTTCTGAACCAACGGTGTGCTCCATTCGG 1539
 |||||

QY 303 LeuGlySerTyrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
 |||||
 DB 1540 CTGGGTTCTCGGTTCCGAAATCGGTATCGTCAACAGTCGTCCTATGCTATGTGCGA 1599
 |||||

QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
 |||||
 DB 1600 CCTAATGTCATCACTTAGCTCTGCCCTTAACATGTTGTTCAAGCTGGAGTGAGTCAATA 1659
 |||||

QY 343 ProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 |||||
 DB 1660 CCNNNNNNNNNNNNNNNNNNCTGCTTGCATCTTACTCTAGAGTGACATCAGCC 1713
 |||||

QY 363 LeuAlaLysAspIleValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
 |||||
 DB 1720 CTAGCTAAAGCCAGAGTCTACACATGTTGCCCTGGAGTGAATGCCATTCATCAGGT 1779
 |||||

QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
 |||||
 DB 1780 TTACAGCTTGCACAAAGATGCCAATACGCTTTTGGANNNNNNNAACGTCAGAGATGCCAG 1839
 |||||

QY 403 MetCysLeuArgSerAspProLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
 |||||
 DB 1840 ATGTGCTCCGGTCTGATCCTTTGCTGAAGTGATGGAGCGCATGGCTAATCCNNNNNG 1899
 |||||

QY 423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer 442
 |||||
 DB 1900 CGCGGGTGTTCATAGTAGAGCTGGAGCAAACTGTGGAGGGTATTATATCATCTAGT 1959
 |||||

QY 443 AspIlePheLysPheLeuLeuSerLeu 451
 |||||
 DB 1960 GATATTTTCAAGTCTTCTGCTGAGCTTG 1986
 |||||

RESULT 2
 CD434895 873 bp mRNA linear EST 03-JUN-2003
 LOCUS E010330A04.b Endosperm_3 Zea mays cDNA, mRNA sequence.
 DEFINITION CD434895
 ACCESSION CD434895
 VERSION CD434895.1 GI:31350538
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 ORGANISM SM

Db	618	ACTTATTATTCGGGGAGCTAGAAACTCATGCTCGAACTTTGACAGAAGATCAGCTTGAAC	677							
Qy	201	xHisThrIleSerAlaTrrpysgluhalalysArgGlnThrAsnGlyArgAsnAspSerG	221							
Db	678	ACACATATATCTGCGATGGAAGAGGCTTAAGCGGCAACCTTTGGAAGAAATGATGCTCA	737							
Qy	221	nTrrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAl	241							
Db	738	GTGGCGAGCACATCAGCACTAGTGCATGCCACCCCTTATGACTCCTTGAGGGACATTGC	797							
Qy	241	aValIysLeuLeuGlnAnsnGlyIleSerThrValProValIleTyrSerSerSerAs	261							
Db	798	AGGTAAGAGCTTTGCAAAATGACATTTCTACAGTGCAGTCCATTATTATTCATCATCAGA	857							
Qy	261	pGlySerPheProGln	266							
Db	858	TGGATCATTCCTCCCTCAG	873							
RESULT 3	CA211312	666 bp	mRNA	linear	EST 25-SEP-2003					
LOCUS	SCRLAD1098B08.g	Ad1	Saccharum officinarum	cdna	clone	SCRLAD1098B08				
DEFINITION	5', mRNA sequence.									
ACCESSION	CA211312									
VERSION	CA211312.1	GI:35255493								
KEYWORDS	EST.									
SOURCE	Saccharum officinarum									
ORGANISM	Saccharum officinarum									
REFERENCE	Eukariota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.									
AUTHORS	1 (bases 1 to 666)									
TITLE	Vettore,A.L., da Silva,F.R., Kemper,E.B. and Arruda,P.									
JOURNAL	The libraries that made SUCST									
COMMENT	Genet. Mol. Biol. 24 (1-4), 1-7 (2001) Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br Plate: 098 row: B column: 08 Seq primer: T7 Promoter Primer. Location/Qualifiers 1..666 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCRLAD1098B08" /lab_host="DH103" /clone_lib="AD1" /note="Organ: seedlings inoculated with Gluconacetobacter diazotrophicans; Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from seedlings inoculated with Gluconacetobacter (diazotrophicans). cDNA was prepared from polyA+ mRNA using super-script plasmid system kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose Chi-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucst.lad.ic.unicamp.br/public "									
FEATURES	source									
ORIGIN										
Alignment Scores:										
Pred. No.:	2.43e-117	Length:	666							
Score:	1090.00	Matches:	211							
Percent Similarity:	95.52%	Conservative:	2							

Best Local Similarity: 94.62% Mismatches: 8
 Query Match: 46.54% Indels: 2
 DB: 14 Gaps: 0

US-09-857-525C-2 (1-45) x CA211312 (1-666)

```

QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
DB 1 CAATTTGCGGATGTCCTCTGTCGAGGCTGCCCATCTGATTTTCAGGCTATTTGCGACCTG 60
QY 24 SerProGlyIleHisGluThrLysPhePheValAspGlyGluTrpArgHisAspGluArg 43
DB 51 TCTCCAGGATCCAGAGTACAGATTTTGTGGACGGAGTGGCGCATGATGAGCGC 120
QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuThrArgGluThr 63
DB 121 CAACCTACTATATCCGGGAGTGTGGTATAGTTAACAACATCTACTTGCACAGAAATTT 180
QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
DB 181 AACCAAAATAAACGCTTTTAAAGTCCAGCAAAATTTGGAAGCAGGATGAACATGATGTG 240
QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
DB 241 GATATGAAATTTTCAAGTACGGTACGTTGTGCGATGGTACTGTTTCGAGGACT 300
QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluThrLeuAsn 123
DB 301 CCGAGAGTTTCAGAGGCTGCATACAAATCTCTAGTGCGCGCTTCTGAATATCTGAAT 360
QY 124 LeuHisThrCysTrpAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
DB 361 TTGCATACATGATGATTTTCTCCGAGTTCAGGCAAGTTATGCTCTGACATTAAT 420
QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
DB 421 TTACCTGTGAAGCAATCTTTTTCATATCTTACATGAACAGGGGATTCCTGTAGCTCT 480
QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 183
DB 481 TGGGACTCAATTCAGAGGTCAATTTGTGTCTCTGAGCCCATGGATTCATATCTATA 540
QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThr 203
DB 541 TCGGGAGCTAGAACTCATGCTCGACCTGACAGACAGACAGCTTGAACACACACT 600
QY 204 TieserAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln-TrpA 223
DB 601 ATATCGATGGAAGAGGCTAACCGGCAAACTTTTATCGAGAAATGATGGTCAATTGGC 660
QY 223 r3Pro 224
DB 661 GACCG 665

```

RESULT 4
 AW744961
 LOCUS
 DEFINITION LG1_385.D1.0.bl_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA sequence.
 ACCESSION AW744961
 VERSION AW744961.1 GI:7658699
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 644)
 Cordonnier-Pratt M.-M., Gingle, A., Marsala, C. and Pratt, L. H.
 An EST database from Sorghum: light-grown seedlings
 TITLE Unpublished (2000)
 JOURNAL Contact: Cordonnier-Pratt MM
 COMMENT Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
 Seq primer: JEN REV
 High quality sequence stop: 621
 POLYA=No.

FEATURES

Location/Qualifiers
 source
 1..644

/organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Light Grown 1 (LG1)"
 /note="Organ: 10- to 14-day-old light-grown (greenhouse) seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN

Alignment Scores:
 Pred. No.: 5,16e-117 Length: 644
 Score: 1087.00 Matches: 285
 Percent Similarity: 98.11% Conservative: 3
 Best Local Similarity: 96.70% Mismatches: 4
 Query Match: 46.41% Indels: 0
 DB: 10 Gaps: 0
 US-09-857-525C-2 (1-451) x AW744961 (1-644)
 QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 DB 4 GAACATTTGCCGATGCTCTCTCGAAGGCTGCCCATCTGATTTTCAGGCTATTTSCAGC 63
 QY 23 LeuSerProGlyIleHisGluThrLysPheValAspGlyGluTrpArgHisAspGlu 42
 DB 64 CTGCTCCAGGAGTTCAGAGTACAAAGTTCTTTTGGATGGGAGTGGCGCATATGAG 123
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuThrArgGlu 62
 DB 124 CCGCAACCTACTATATCTGGGAGGTTTGAATAGTTAACACATTACTTGCACAGGGAA 183
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 DB 184 TTTAACCAACTAAACGCCCTTATTAAGTCCAGACACACCTCGAAGCAGGATGAACATGGAT 243
 QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 DB 244 GTGGATAATGAAATTTTCAAGTACGGTTAGTTGTGCGATGGTACCTTCGGAGGTT 303
 QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluThrLeu 122
 DB 304 ACTCCGAGAGTTTCAGAGGCTGCATACAAATCTCTAGTGCGCGCTTCTGAATATCTG 363
 QY 123 AsnLeuHisThrCysTrpAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
 DB 364 AATTTGCATACATGCTATGATTTACTCCGGATTCGGGCAAGGTTATTTGCTCTGACAT 423
 QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 DB 424 AATTTACTGTGAAGCAATCTTTTCATATCTGCAATGACAGGGATTCCTGTAGCTCT 483
 QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 DB 484 CTCTGGGACTCATTTAGAGGTCAATTTGTGTCTTCTGAGCCCATTTGATTTTCATCT 543
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
 DB 544 ATATTGGGAGCTAGAAACTCATGGCTCGAATTTGACAGAGACGCTTGAAACACAT 603

Best Local Similarity: 94.62% Mismatches: 8
 Query Match: 46.54% Indels: 2
 DB: 14 Gaps: 0

US-09-857-525C-2 (1-45) x CA211312 (1-666)

```

QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
DB 1 CAATTTGCGGATGTCCTCTGTCGAGGCTGCCCATCTGATTTTCAGGCTATTTGCGACCTG 60
QY 24 SerProGlyIleHisGluThrLysPhePheValAspGlyGluTrpArgHisAspGluArg 43
DB 51 TCTCCAGGATCCAGAGTACAGATTTTGTGGACGGAGTGGCGCATGATGAGCGC 120
QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuThrArgGluThr 63
DB 121 CAACCTACTATATCCGGGAGTGTGGTATAGTTAACAACATCTACTTGCACAGAAATTT 180
QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
DB 181 AACCAAAATAAACGCTTTTAAAGTCCAGCAAAATTTGGAAGCAGGATGAACATGATGTG 240
QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
DB 241 GATATGAAATTTTCAAGTACGGTACGTTGTGCGATGGTACTGTTTCGAGGACT 300
QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluThrLeuAsn 123
DB 301 CCGAGAGTTTCAGAGGCTGCATACAAATCTCTAGTGCGCGCTTCTGAATATCTGAAT 360
QY 124 LeuHisThrCysTrpAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
DB 361 TTGCATACATGATGATTTTCTCCGAGTTCAGGCAAGTTATGCTCTGACATTAAT 420
QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
DB 421 TTACCTGTGAAGCAATCTTTTTCATATCTTACATGAACAGGGGATTCCTGTAGCTCT 480
QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 183
DB 481 TGGGACTCAATTCAGAGGTCAATTTGTGTCTCTGAGCCCATGGATTCATATCTATA 540
QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThr 203
DB 541 TCGGGAGCTAGAACTCATGCTCGACCTGACAGACAGACAGCTTGAACACACACT 600
QY 204 TieserAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln-TrpA 223
DB 601 ATATCGATGGAAGAGGCTAACCGGCAAACTTTTATCGAGAAATGATGGTCAATTGGC 660
QY 223 r3Pro 224
DB 661 GACCG 665

```

RESULT 4
 AW744961
 LOCUS
 DEFINITION LG1_385.D1.0.bl_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA sequence.
 ACCESSION AW744961
 VERSION AW744961.1 GI:7658699
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 644)
 Cordonnier-Pratt M.-M., Gingle, A., Marsala, C. and Pratt, L. H.
 An EST database from Sorghum: light-grown seedlings
 TITLE Unpublished (2000)
 JOURNAL Contact: Cordonnier-Pratt MM
 COMMENT Laboratory for Genomics and Bioinformatics

QY 203 ThrileSerAlaTrpLysGluAlaLysArgGlnThr 214
 |||||
 Db 604 ACGATATCTGCATGAAAGAAAGCTTAAGCGGCAACT 639

RESULT 5
 CD878389 738 bp mRNA linear EST 11-JUL-2003
 LOCUS AZ04.102K03F010930 AZ04 Triticum aestivum cDNA clone AZ04102K03,
 mRNA sequence.

ACCESSION CD878389
 VERSION CD878389.1 GI:32562205
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 738)
 AUTHORS Genoplante.
 TITLE Genoplante, a major partnership french program in plant genomics
 JOURNAL Unpublished (2003)
 COMMENT Contact: Genoplante
 Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' {http://www.genoplante.com
 and http://genoplante-info.inbioigen.fr}.

FEATURES
 source 1..738
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="recital"
 /db_xref="taxon:4565"
 /clone="AZ04102K03"
 /tissue_type="root"
 /clone_lib="AZ04"

ORIGIN
 Alignment Scores:
 Pred. No.: 8.71e-117 Length: 738
 Score: 1086.00 Matches: 209
 Percent Similarity: 91.80% Conservative: 15
 Best Local Similarity: 85.66% Mismatches: 18
 Query Match: 46.37% Indels: 2
 DB: 14 Gaps: 1

US-09-857-525C-2 (1-451) x CD878389 (1-738)

QY 75 ProGlySerArgMetAsnMetAspValAspAsnGluAsnPheGlnArgThrValThrLeu 94
 |||||
 Db 12 CCTGGGAGCAGG-----ATGGATGTGGACAATGATGTTTCAACGAATGGGTTCGTTG 65

QY 95 SexAspGlyThrValSerGluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSer 114
 |||||
 Db 66 TTGGATGTGTCCTTCAGGAAGTTCCTCCAGAGATTTTCAGAGGCTGCTATACAGATCTCT 125

QY 115 ArgCysArgValSerGluThrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSer 134
 |||||
 Db 126 AGGTGTCTGTGTGTGAGTATCTGAATGCGCATACAGGCCATGACCTACTACCAATCTCT 185

QY 135 GlyLysValIleAlaLeuAspIleAsnLeuProValLysGlnSerPheHisIleLeuHis 154
 |||||
 Db 186 GGAAGATCATGCTCTGGACATTAATTTACCTGTGAAGCAACTTTTCATATCTCTCAT 245

QY 155 GluGlnGlyIleProValAlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeu 174
 |||||
 Db 246 GAACAGGGGATCTCTGTGGCTCTCTGTGGATTCATTCAGGGGTGAGTTGTGGCCTT 305

QY 175 LeuSerProLeuAspPheLeuLeuLeuArgGluLeuGluThrHisGlySerAsnLeu 194
 |||||

Db 306 CTGAGCCACTGGATTTTATTAATTTATATTTGAGAGAGCTGGAAACGCATGGCTCAACCTG 365

QY 195 ThrGlnGlnLeuGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThr 214
 |||||
 Db 366 ACAGAGAAACAGCTTGAACAACACACATATATCTCGTGGAAAGAGCTTAAGCGGCAACT 425

QY 215 AsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyr 234
 |||||
 Db 426 TATGGAGAATGATGGACAATCTAGATCAATCAGCATCTAGTGCATGCCACCCCTTAT 485

QY 235 GluSerLeuArgAspIleAlaValLysLeuLeuGlnAsnGlyIleSerThrValProVal 254
 |||||
 Db 486 GAATCCTTGGGGTATTGCCATGAAATATCTTGAACCTGGTATTCTACAGTCAACATC 545

QY 255 IleTyrSerSerSerSerAspGlySerPheProGlnGlnLeuHisLeuAlaSerLeuSer 274
 |||||
 Db 546 ATCTATTCATCGTCATCAGATGGATCGTTTCGACGCTGTGCATCTTGATCCCTTCA 605

QY 275 GlyIleLeuLysCysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeu 294
 |||||
 Db 606 GGAATTTTGAATGTATCTGTAGATACCTTAAGAACTCCACTGTGTAGTTTGGCATTTCTA 665

QY 295 AsnGlnProValCysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsn 314
 |||||
 Db 666 AACCAACCAAGTATGCTCAATTCGCTGGGTATACATGGTTCACAAAATTTGGTGAACCAAT 725

QY 315 SerArgProLeu 318
 |||||

Db 726 GGTCAATCCATTTG 737

RESULT 6
 CA500333 737 bp mRNA linear EST 14-NOV-2002
 LOCUS WHE4018_C01_E02ZT Wheat meiotic anther cDNA library Triticum
 DEFINITION aestivum cDNA clone WHE4018_C01_E02, mRNA sequence.

ACCESSION CA500333
 VERSION CA500333.1 GI:24991293
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Authors Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 737)
 Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R.,
 Pham, J., Rausch, C.J., Sutton, T., Woo, J., and Wilson, C.
 The structure and function of the expressed portion of the wheat
 genomes - Meiotic anther cDNA library
 Unpublished (2002)
 Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818

Seq primer: 5' primer.
 Location/Qualifiers
 1..737
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE4018_C01_E02"
 /tissue_type="Anther"
 /dev_stage="Meiotic stages pre-meiosis-metaphase I"
 /lab_host="E. coli DH10B"
 /clone_lib="Wheat meiotic anther cDNA library"
 /note="Vector: pSPOR1; Site1: SalI; Site2: NotI; Plants
 were grown in a glasshouse. Anther meiotic stage was
 determined by removing anthers from individual primary

florists. One author was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two authors were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared, cDNA synthesized, and directionally ligated into pSPOR1 by Jim Sutton in the p Langridge lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

Alignment Scores:
Pred. No.: 3.42e-114 Length: 737
Score: 1064.00 Matches: 205
Percent Similarity: 92.18% Conservative: 19
Best Local Similarity: 84.36% Mismatches: 17
Query Match: 45.43% Indels: 2
DB: 14 Gaps: 1

US-09-857-525C-2 (1-451) x CA500333 (1-737)

QY 51 PheGlyIleValAsnThrLeuThrLeuThrArgGluTyrAsnGlnIleAsnThrLeuSer 70
DB 13 TATGGGGTGGTAAACACCTTGTACTTGACAGGGAAATTGACCATAAATCTACTG 72
QY 71 SerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPheGlnArg 90
DB 73 AGCCCACTACACCTGGGAGCAGG-----ATGGATGTGGACAATGATAGTTTCAACGA 126
QY 91 ThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGluAla 110
DB 127 ATGGGTTCCTGTTGGATGGTGCCTTCTCAGGAAGGTTCTCCAAGAATTTCAAGGCTCT 186
QY 111 IleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyrAspLeu 130
DB 187 ATACAGATCTCTAGGTGCTGTTGCTGAGTATCTGAATGGCATACAGGCTACGACCTA 246
QY 131 LeuProAspSerGlyValIleAlaLeuAspIleAsnLeuProValIleGlnSerPhe 150
DB 247 CTACCAGATCTGGAAAGCTCATGTCTGGACATTAATTACCTGTGAGCAATCTTTC 306
QY 151 HisIleLeuHisGluGlnGlyIleProValAlaProLeuThrAspSerPheArgGlyGln 170
DB 307 CATATCTCCATGAACAGGGGATTCCTGTGCTCTCTGTGGGATTCATTCAGGGGTCCAG 366
QY 171 PheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGluThrHis 190
DB 367 TTTGTGGCTTCTGAGCCCACTGGATTTTACTTATATTCAGAGAGCTGGAACGGAT 426
QY 191 GlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSerAlaThrIleGluAla 210
DB 427 GGCTCAAACTCCACAGAGGAACAGCTTGAACACACACTATATCTGCGTGGAAAGAGCT 486
QY 211 LysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLeuValHis 230
DB 487 AAGCGGCAAACTTATGGAAGAAATGATGGAACAACCTAGATCAATCAGCATCTAGTGCT 546
QY 231 AlaThrProTyrGluSerLeuArgAspIleAlaValIleLeuLeuGlnAsnGlyIleSer 250
DB 547 GCCACCCCTTATGAATCCTTGGAGGGTATTGCCATGAATAACTTGAACCTGATTTCT 606
QY 251 ThrValProValIleTyrSerSerSerSerAspGlySerPheProGlnLeuLeuHisLeu 270
DB 607 ACAGTACCAATCATCTATTATCGTCATCAGATGATGATGCTTTCGACGCTGTGTCATCT 666
QY 271 AlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSerThrGlyAsn 290
DB 667 GCATCCCTTTCAGGAATTTGAATGATACCTGATAGTACTTTAAGACTCCACTGGTNGT 726
QY 291 LeuProIle 293
|||||

DB 727 TTGCCGATT 735

RESULT 7

CD903099
LOCUS CD903099 712 bp mRNA linear EST 14-JUL-2003
DEFINITION G356.109C03F010918 G356 Triticum aestivum cDNA clone G356109C03,
mRNA sequence.
ACCESSION CD903099
VERSION CD903099.1 GI:32677427
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum

REFERENCE Sukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.

AUTHORS 1 (bases 1 to 712)

TITLE Genoplante, a major partnership french program in plant genomics

JOURNAL Unpublished (2003)

COMMENT Contact: Genoplante

Genoplante

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french

plant genomics programme "Genoplante" (<http://www.genoplante.com>

and <http://genoplante-info.infobiogen.fr>).

FEATURES

source

1..712

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="recital"

/db_xref="taxon:4565"

/clone="G356109C03"

/tissue type="grain (356 degrees per day after

pollination)"

/clone_lib="G356"

ORIGIN

Alignment Scores:
Pred. No.: 1.44e-112 Length: 712
Score: 1050.00 Matches: 203
Percent Similarity: 94.30% Conservative: 12
Best Local Similarity: 89.04% Mismatches: 12
Query Match: 44.83% Indels: 1
DB: 14 Gaps: 0

US-09-857-525C-2 (1-451) x CD903099 (1-712)

QY 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
DB 12 ACTATATCTCGGTGAAAGAGGCTAAGCGCAACTTATGGAAGAAATGATGGCAACT 71
QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
DB 72 AGATCAATACAGCATCTAGTGCATGCCACCTTATGATCTCTGAGGGGTATGCCATG 131
QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
DB 132 AAAATACCTCGAAACTGGCATCTTACAGTCCCAATCATCTATTCATCGATCGA 191
QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
DB 192 TCGTTTCCGACGCTGTGCATCTTGCATCCCTTTTCAGGAATTTGAAATGATCTCTAGA 251
QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
DB 252 TACTTCAAGAACTCCACTGGTAGTTTCCCGATTCTAAACCAACAGATGCTCAATTCCG 311
QY 303 LeuGlySerThrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
DB 312 CTGGGTACCTGGGTTCAAAAAATTTGGTGAACCAAAATGGTTCATCCATTTGGCTATGTCGG 371

QY 323 ProAenAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerile 342
 Db 372 CCTAATACATCTCTAGCTCTGCCCTTAACCTGTGTGGTTCAAGCTGGAGTAGTTCATA 431
 QY 343 ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 Db 432 CCCATTGTGGATGATACGACTCGCTGATCGACACATACACTCCAGAGTGACATCACAGCT 491
 QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
 Db 492 CTAGCGAAAGCAAGGCTACACCCATATCCGCTAGATGAGATGACCATTCATCAGGCC 551
 QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
 Db 552 TTGCAGCTCGGCAAGACGCGAATTCACCTTTTGACCTTTTCATGGTCAAGATGACG 611
 QY 403 MetCysLeuArgSerAspProLeuLeu-LysValMetGluArgLeuAlaAsnProGlyVa 422
 Db 612 ATGTGTCTCCAGCTGACCCCTTTCGTAAGGTTATGAGAGATTGGCTAATCTCTGGGGT 671
 QY 422 LArgAtcValPheIleValGlu 429
 Db 672 GCGTCGCGGTTCATCGTGAA 693

RESULT 8
 CA262684 683 bp mRNA linear EST 26-SEP-2003
 LOCUS SCPIILB2024A07.g LB2 Saccharum officinarum cDNA clone SCPIILB2024A07
 DEFINITION S', mRNA sequence.

ACCESSION CA262684
 VERSION CA262684.1 GI:35954887
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum

REFERENCE
 AUTHORS Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 TITLE The libraries that made SUCEST
 JOURNAL Genet. Mol. Biol. 24 (1-4): 1-7 (2001)

COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parvadaeunicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccen.br/fcav.unesp.br

Plate: 024 row: A column: 07
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1. 683

FEATURES
 source
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCPIILB2024A07"
 /lab_host="DH10B"
 /clone_lib="JB2"
 /notes="Organ: Lateral buds from plants adult plants
 growing in greenhouse; Vector: pSport1; Site 1: SalI;
 Site 2: NotI; An unidirectional cDNA library generated
 from [lateral buds from plants adult plants growing in
 greenhouse]. cDNA was prepared from polyA+ mRNA using
 Superscript Plasmid System Kit (Invitrogen). The
 double-strand cDNAs were fractionated in a sepharose
 CL-2B 40cm-columns and fragments sizing between 0.8 and
 1.5 Kb were directionally cloned into the vector. Details
 of each source of RNA and library construction can be
 obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
 Pred. No.: 5,18e-112 Length: 683
 Score: 1045.00 Matches: 205
 Percent Similarity: 94.52% Conservative: 2
 Best Local Similarity: 93.61% Mismatches: 12
 Query Match: 44.62% Indels: 1
 DB: 14 Gaps: 0

US-09-857-525C-2 (1-451) x CA262684 (1-683)

QY 3 GluHisLeuProMetSerProIleGlyCysProThrValPheGlnAlaIleCysSer 22
 Db 26 GAACATTTGTGATGTCTCTCTGCTGT-AGTTGCCCACTGTATTTCAGGCTATTTCAGC 84
 QY 23 LeuSerProGlyIleHisGluTyrLysPheValAspGlyGluTyrArgHisAspGlu 42
 Db 85 CTGTCTCCAGGGAATTCAGAGTCAAGGCTTTTGTGGACGGGAGTGGCGGCATGATGAG 144
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 Db 145 CGCCAACCTACTATATCCGGGAGTTTGGTATAGTTAAACATTCATTGTTGACAGAGAA 204
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 Db 205 TTTAAACCAATATAAAGCGCTTATTAACTCCAAAGCACAAATTGGAAGCAGGATGAACATGGAT 264
 QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 Db 265 GTGGATAATGAAAAATTTTCAACGCTACGGTTACGTTGTCGGATGCTACTGTTTCGGAAAGCT 324
 QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 Db 325 ACTCGAGAGTTTCAGAGCTGCAATACATCTAGTGCCTGCGGCTTCTGAAATATCTG 384
 QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
 Db 385 AATTGGCATACATCTATGATTACTCCGGATTTCAGGCAAGGTTATTGCTCGGACATT 444
 QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 Db 445 AATTACCTGTGAAGCAATCTTTTCATATCTACATGAAAGGAGGATTCCTGTAGCTCCT 504
 QY 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 Db 505 CTGTGGGACTCATTCACAGGTCATTTGTTGGTCTTCTGAGCCCATTCGATTTCATCTT 564
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
 Db 565 ATATTGGCGGAGCTAGAAACTCATGGCTCGAACTCGACAGAAAGCAGCTTTGAAACACAC 624
 QY 203 ThrIleSerAlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln 221
 Db 625 ACTATATCTCATGGAAGAGGCTAAGCGGCACACTTATCGAAGAAATGATGGTCAG 681

RESULT 9

BU098591

LOCUS

DEFINITION

946136F11.y1 946 - tassal primordium prepared by Schmidt lab Zea

ACCESION

BU098591

VERSION

BU098591.1

KEYWORDS

EST.

SOURCE

Zea mays

ORGANISM

Zea mays

REFERENCE

AUTHORS

TITLE

JOURNAL

BU098591 601 bp mRNA linear EST 29-AUG-2002
 946136F11.y1 946 - tassal primordium prepared by Schmidt lab Zea

ACCESION BU098591
 VERSION BU098591.1
 KEYWORDS EST.
 SOURCE Zea mays

ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 601)
 AUTHORS Walbot, V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 University

JOURNAL Unpublished (1999)

COMMENT

Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946136 row: F column: 11.
Location/Qualifiers

FEATURES
source

1..601
/organism="Zea mays"
/mol_type="mRNA"
/cultiivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/clone_lib="946 - tassel primordium prepared by Schmidt lab"
/note="Organ: tassels; Vector: HybridZAP; Site 1: EcoRI; Site 2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybridZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

ORIGIN

Alignment Scores:
Pred. No.: 6.25e-109 Length: 601
Score: 1018.00 Matches: 193
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 43.47% Indels: 0
DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x BU098591 (1-601)

Qy 3 GluHisLeuProMetSerProileGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 21 GAAACATTTSCGAGTCTCTCTATCGAGGCTGCCCATGTATTTTCAGGCTATTTGCAGT 80
Qy 23 LeuSerProGlyIleHisGluTrpLysPheValAspGlyGluTrpArgHisAspGlu 42
Db 81 CTGCTCTCTGGGATTCACGAGTACAAATTCCTTTGTGGACGGGAATGGCGGCATGATGAG 140
Qy 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrlleuThrArgGlu 62
Db 141 CGTCAACCTACCATATCTGGGGAGTTGGCATAGTTAAACACACTTTTCTTGACAAAGGAA 200
Qy 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
Db 201 TATAACCAATAAATACACTTATCAAGTCCAGCACACCTGGAGCAGGATGAACATGGAT 260
Qy 83 ValAspAsnGluAsnPheGluArgThrValThrLeuSerAspGlyThrValSerGluGly 102
Db 261 GTGGATATGAAAATTTTCAACGTACGGTTACGTTGTGATGATGGTACCGTTTCAGAAAGT 320
Qy 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrlleu 122
Db 321 ACTCTGAGAGTTTCAGAGCTCAATACAAATATCTAGGTGCGTGTCTGAATATCTG 380
Qy 123 AsnLeuHisThrCysTyAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 381 AATTTGCATACATGCTATGATTTTCTCCAGATTTCTGGCAAGTTTATTCCTCAGATTT 440
Qy 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db 441 AATTTTACCTGTGAGCAATCATTCATATTCATATTCCTCCATGAACAGGGGATTCCTAGCTCT 500
Qy 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 501 CTCTGGGACTCATTCAGAGGTCAATTTGTGTGCTCTCTTAGCCCAATGGATTTCTATCT 560

Qy 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThr 195
Db 561 AATTTGCGGAGCTAGAAACTCATGGCTCGAATTGACA 599

RESULT 10

BE598304

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Cordnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions

below Phred quality 16. The threshold for highest quality sequence

is 20.

Seq primer: JEN REV

High quality sequence stop: 546

POLYA=No.

FEATURES
source

Location/Qualifiers

1..595

/organism="Sorghum bicolor"

/mol_type="mRNA"

/db_xref="taxon:4558"

/clone_lib="Pathogen induced 1 (PII)"

/note="Organ: Anthracnose-infected leaves from

two-week-old sorghum plants 48 hr after inoculation;

Vector: pBluescript II from Lambda Zap II; Site 1: XhoI;

Site 2: EcoRI; Two-week-old sorghum plants (BTK 623

cultivar) were infected with pathogen (isolate PRM421 of

Colletotrichum graminicola, which is a sorghum isolate).

RNA was prepared from infected leaves harvested from 45

seedlings 48 hours after inoculation. Note: young

seedlings (2 weeks old) exhibit juvenile resistant

reaction, which is an incompatible interaction. As they

grow older (4 weeks or older), plants resume susceptibility

to anthracnose disease. The library was made from Poly-A

RNA in the cloning vector lambda Zap II. Clones to be

sequenced were prepared by mass excision. WARNING: While

most or all ESTs are expected to derive from the host

plant, no effort was made to eliminate ESTs deriving from

the pathogen."

ORIGIN

Alignment Scores:

Pred. No.: 1.07e-106 Length: 595

Score: 999.00 Matches: 190

Percent Similarity: 96.95% Conservative: 1

Best Local Similarity: 96.45% Mismatches: 6

Query Match: 42.66% Indels: 0

DB: 10 Gaps: 0

US-09-857-525C-2 (1-451) x BE598304 (1-595)

Qy 119 SerGluTyrlleuAsnLeuHisThrCysTyAspLeuLeuProAspSerGlyLysValIle 138

Db 3 ACGAGGTATCTGAATTTGTCATACATGCTATGATTTACTCCCGGATTCGGGCAAGGTATT 62

```

Qy 139 AlaLeuAspIleAsnLeuProVallyLeuGlnSerPheHisIleLeuHisGluGlnGlyLe 158
Db 63 GCTCTGGACATTAATTTACCTGTGAAGCAATCTTTTCATATTCACATGACAGGGGATT 122
Qy 159 ProValAlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeu 178
Db 123 CTTGTAGTCTCTCTGGGACTCATTTAGAGGTCAATTTGTGTCTCTCTGAGCCATTG 182
Qy 179 AspPheIleLeuLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGln 198
Db 183 GATTTTCATCTTATATTGGGGAGCTAGAAACTCATGGCTCGAACTTGCAGAGAGACAG 242
Qy 199 LeuGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsn 218
Db 243 CTTGAACACATACGATATCTGCATGGAAGAGGCTAAGCGGCAACTTATGAGGAAT 302
Qy 219 AspSerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGlnSerLeuArg 238
Db 303 GATGGTCAATGGCGCACCATCAGCATCTAGTGCATGCTACCCCTTATGAGTCTTTGAGG 362
Qy 239 AspIleAlaVallyLeuLeuGlnAsnGlyLeuSerThrValProValIleTyrSerSer 258
Db 363 GACATTCGAGTAAACTTTTGCAAAATGGCATTTCTACAGTGCAGTTATTTATTCATCA 422
Qy 259 SerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyLeuLeuLys 278
Db 423 TCATCGATGGATCATTCCTGGCAGTTGTTGCATCTTGCATCCCTCTCTGGAAATTTGAAA 482
Qy 279 CysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProVal 298
Db 483 TGTATTTGTAGATATTTCAAAAACCTCGACTGTGAATTTGCCTATTCTGAAACCAACAGTG 542
Qy 299 CysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSer 315
Db 543 TGCTCATTCACCTGGGTTCTGGGTTCCGAAATTTGGTATCTTCAACAGT 593

```

```

RESULT 11
CA218557 577 bp mRNA linear EST 25-SEP-2003
LOCUS SCUFADIC09B02.g Adl Saccharum officinarum cDNA clone SCUFADIC09B02
5', mRNA sequence.
CA218557
CA218557.1 GI:35269977
EST.
Saccharum officinarum
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: C09 row: B column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..577
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCUFADIC09B02"
/lab_host="DH10B"

```

FEATURES

```

source
BE123279
LOCUS BE123279
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE123279
VERSION BE123279.1 GI:8516554
KEYWORDS EST.
SOURCE Zea mays

```

```

/clone_lib="AD1"
/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Gluconacetobacter
diazotrophicans]. cDNA was prepared from polyA+ mRNA using
SuperScript plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"

```

ORIGIN

```

Alignment Scores: 3.49e-103 Length: 577
Pred. No.: 969.00 Matches: 184
Score: 97.38% Conservative: 2
Percent Similarity: 96.34% Mismatches: 5
Best Local Similarity: 41.37% Indels: 0
Query Match: 14 Gaps: 0
DB: 14

US-09-857-525C-2 (1-451) x CA218557 (1-577)

Qy 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
Db 1 CATTTGCCGATCTCTCTGTGGAAGGCTGCCCACTGTATTTTCAGGCTATTTGCAGCTG 60
Qy 24 SerProGlyIleHisGluTyrLysPhePheValAspGlyGluTrpArgHisAspGluArg 43
Db 61 TCTCCAGGGATCCACGAGTACAGTCTTTTGTGGACGGAGTGGCGGCATGATGAGCGC 120
Qy 44 GlnProTrpIleSerGlyGluPheGlyLeuValAsnThrLeuTyrLeuThrArgGluTyr 53
Db 121 CACCTTACTATATCCGGGGAGTTTGGTATAGTTAAACACATTTGTACTTGCACAGAGAAATTT 180
Qy 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
Db 181 AACCAATAAACGCTTATTAAAGTCCAGCACAAATTCGAAGCAGAGATGACATGATGTG 240
Qy 84 AspAsnGlnAsnThrGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
Db 241 GATAATGAAATTTTCAACGTCAGGTTACGTTGTGGATGGTACTGTTCGGAAGGTAAT 300
Qy 104 LeuArgValSerGlnAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsn 123
Db 301 CCAGAGATTTCAGAGGCTGCATACAAATCTTAGGTGCCCGCTTCTGAATATCTGAAT 360
Qy 124 LeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
Db 361 TTGCATACATGCTATGATTTACTCCCGGATTCAGGCAAGGTTATTTGCTTGGACATTAAT 420
Qy 144 LeuProVallyGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
Db 421 TTACCTGTGAAGCAATCTTTTCATATTTCTACATGAAACAGGGGATTCCTGTAGCTCTCTG 480
Qy 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
Db 481 TGGGATCATTCAGAGTCAATTTGTGTGTTCTTAGGCCCATTTGGAATTTCTACTATTATA 540
Qy 184 LeuArgGluLeuGluThrHisGlySerAsnLeu 194
Db 541 TTGGGGAGCTAGAAACTCATGCTTGAACCTG 573

```

RESULT 12

```

BE123279 559 bp mRNA linear EST 14-JUN-2003
LOCUS 945040D06.Y1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE123279
VERSION BE123279.1 GI:8516554
KEYWORDS EST.
SOURCE Zea mays

```


ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 559)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945040 row: D column: 36.
FEATURES
Location/Qualifiers
Source 1..559
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/clone_lib="945 - Mixed adult tissues from Walbot lab, same as 707 (SK)"
/note="Organ: tassel, kernal, silk, husk, root, leaf; Vector: pGAD10; Site 1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned. New library number given to library 707 for additional sequencing."
ORIGIN
Alignment Scores:
Pred. No.: 1.29e-102 Length: 559
Score: 964.00 Matches: 186
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.16% Indels: 0
DB: 10 Gaps: 0
US-09-857-525C-2 (1-451) x BE123279 (1-559)
Qy 66 TleAsnThrLeuSerProSerThrProGlySerArgMetAsnMetAspValAspAsn 85
Db 2 ATAAACACCTTATCAAGTCCAGCACACCTGGAGCAGGATGAACATGGATGATAAT 61
Qy 86 GluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArg 105
Db 62 GAAATTTTCAAGTACGGTTACGTTGTCAGATGGTACCGTTTCAGAGGAGTACTCTGAGA 121
Qy 106 ValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHis 125
Db 122 GTTTCAGAGCGCTCAATACAAATATCTAGGTCGTCGTTCTGAATATCTCAATTTGCAT 181
Qy 126 ThrCysTyrAspLeuLeuProAspSerGlyValIleAlaLeuAspIleAsnLeuPro 145
Db 182 ACATGCTATGATTTTACTCCCAATTCCTGGCAAGGTATTCGCCCTAGACATTAATTTACCT 241
Qy 146 ValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTyrAsp 165
Db 242 GTGAAGCAATCTTCATATTTCTCATGAAACAGGGGATTCCTGTAGCTCTCTCTGGAC 301
Qy 166 SerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArg 185
Db 302 TCAFTTCAGAGGTCAATTTGTTGGTCTCTCTAGCCCATTTGGATTTCAATCATATTCGGG 361
Qy 186 GluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSer 205
Db 362 GAGCTAGAAACTCATGGCTCGAATTTGACAGAGAGCAGCTTGAACACACACTATATCT 421

Qy 206 AlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyrArgProGln 225
Db 422 GCATGGAAGAGGCTAAGCGGCAAACTAATGAAGAAATGATGATCAGTGGCGCCGCA 481
Qy 226 GlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeu 245
Db 482 CAGCATCTAGTGCATGCCACCCCTATGAGTCTTTCAGGGACATTCGAGTAAGCTTTTG 541
Qy 246 GlnAsnGlyIleSerThr 251
Db 542 CAAAATGGCATTTCTACA 559
RESULT 13
LOCUS BU499410
DEFINITION 946174G09.yl 946 - tassel primordium prepared by Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION BU499410
VERSION BU499410.1 GI:22819320
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 549)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946174 row: G column: 09.
FEATURES
Location/Qualifiers
Source 1..549
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XL0LR"
/clone_lib="946 - tassel primordium prepared by Schmidt lab"
/note="Organ: tassels; Vector: HybridZAP; Site 1: EcoRI; Site 2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybridZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."
ORIGIN
Alignment Scores:
Pred. No.: 1.63e-102 Length: 549
Score: 963.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.12% Indels: 0
DB: 13 Gaps: 0
US-09-857-525C-2 (1-451) x BU499410 (1-549)
Qy 10 IleGluGlyCysProThrValPheGlnAlaIleCysSerLeuSerProGlyIleHisGlu 29
Db 1 ATCGAAGGCTGCCCATCTGATTTTCAGCTATTTTCAGCTATTTTCAGCTATTTTCAGG 60
Qy 30 TyrLysPhePheValAspGlyGluTyrArgHisAspGluArgGlnProThrIleSerGly 49

Db	409	TTCCGACTTTTCAACGCTCAAGATGCAGATGTGTCTCCGGTCTGACCCCTTGCTGAAG	350
Qy	413	ValMetGluArgLeuAlaAsnProGlyValArgArgValPheIleValIcuAlaGlySer	432
Db	349	GTTATGAGAGATGGCTAATCTCGGGTGCGTCCGGTGTTCATCGTGAGGCTGGCAGC	290
Qy	433	LysArgValGluGlyIleIleSerLeuSerAspIlePheTysPheLeuLeuSer	450
Db	289	AACGAGTGAAGCGTAATATCTCAGCGACGATATTCAGTGTCTCTCAGC	236

Search completed: July 9, 2004, 11:12:52
Job time : 3653 secs

```

FEATURES
source
1. .831
Location/Qualifiers
/organism="Secale cereale"
/mol_type="mRNA"
/cultivar="Puma (winter rye)"
/db_xref="taxon:4550"
/clone="Sc01_08c11"
/tissue_type="leaf, crown"
/dev_stage="seedling three-leaf stage"
/clone_lib="Sc01_AAFc_EC0RC_cold_stressed_winter_rye_seed1
ings"
/notes="Vector: Bluescript SK-/XhoI-EcoRI; Site 1: Eco RI;
Site 2: Xho I; Sampled three-leaf seedlings treated for
one week at 20C, 12 hrs light/day. Library made with
Stratagene UNIZAP XR Kit/Gigapack III Gold Kit. lambda
library is amplified, then mass excised in SOLR cells."

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2004, 08:14:48 ; Search time 673 Seconds
(without alignments)
3266.154 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREKPMSPTEGCTVFQAI.....SKRVEGIISLSDIEKFLLSL 451

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTG.spool/US09857525/runat_07072004_161019_325/app_query.fasta_1.647
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09857525 @CGN 1 1 511 @runat_07072004_161019_325
-NCPU=6 -ICPU=3 -NO MWAP -LARGESUREY -NEG SCORES=0 -WAIT_DSBLCK=100
-LONGLOG -DEV TIMEOUT=130 -WARNY_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

Result No.	Score	Match	length	ID	Description
1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:				
2:	/cgn2_6/ptodata/1/pubpna/PCT_NSW_PUB.seq:				
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:				
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:				
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:				
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:				
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:				
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:				
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:				
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:				
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:				
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:				
13:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:				
14:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:				
15:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:				
16:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:				
17:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:				
18:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:				
19:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	length	ID	Description
1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:				
2:	/cgn2_6/ptodata/1/pubpna/PCT_NSW_PUB.seq:				
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:				
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:				
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:				
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:				
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:				
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:				
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:				
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:				
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:				
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:				
13:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:				
14:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:				
15:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:				
16:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:				
17:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:				
18:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:				
19:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:				

1	2231	95.3	1944	13	US-10-425-114-1948	Sequence 1948, Ap
2	2230	95.2	2228	13	US-10-425-114-27250	Sequence 27250, A
3	2178	93.0	1474	13	US-10-425-114-3817	Sequence 3817, Ap
4	1944	83.0	2980	17	US-10-437-963-2873	Sequence 2873, Ap
5	1802	76.9	1387	13	US-10-425-114-33827	Sequence 33827, A
6	1542	65.8	3077	13	US-10-424-599-44858	Sequence 44858, A
7	1492	63.7	1200	13	US-10-425-114-6702	Sequence 6702, Ap
8	1487	63.5	2315	13	US-10-424-599-79689	Sequence 79689, A
9	1325	56.6	1507	13	US-10-425-114-29527	Sequence 29527, A
10	924.5	39.5	1679	13	US-10-424-599-105	Sequence 105, App
11	921	39.3	1225	13	US-10-424-599-101	Sequence 101, App
12	908.5	38.8	6747	17	US-10-437-963-90730	Sequence 90730, A
13	821.5	35.1	1402	17	US-10-437-963-66100	Sequence 66100, A
14	766	32.7	1112	17	US-10-437-963-22833	Sequence 22833, A
15	761	32.5	1533	13	US-10-425-114-8857	Sequence 8857, Ap
16	728.5	31.1	1127	13	US-10-434-599-100	Sequence 100, App
17	679.5	29.0	1001	13	US-10-435-114-10694	Sequence 10694, A
18	580.5	24.8	954	13	US-10-435-114-2008	Sequence 2008, Ap
19	460	19.6	282	9	US-09-294-0938-135	Sequence 135, App
20	442	18.9	365	12	US-09-732-627A-260	Sequence 260, App
21	425.5	18.2	2362	13	US-10-425-114-31135	Sequence 31135, A
22	419	17.9	1140	13	US-10-424-599-44867	Sequence 44867, A
23	373	15.9	1435	10	US-09-769-970-14	Sequence 14, Appl
24	367	15.7	1578	16	US-10-305-720-1113	Sequence 1113, Ap
25	367	15.7	1691	9	US-09-925-297-2	Sequence 2, Appl
26	359	15.3	721	13	US-10-425-114-30669	Sequence 30669, A
27	340	14.5	1647	9	US-09-826-581-5	Sequence 5, Appl
28	340	14.5	1647	17	US-10-705-137-5	Sequence 5, Appl
29	340	14.5	2781	17	US-10-473-670-30	Sequence 30, Appl
30	335.5	14.3	1873	13	US-09-950-022-7	Sequence 7, Appl
31	335.5	14.3	1873	13	US-09-950-022-9	Sequence 9, Appl
32	334.5	14.3	1873	13	US-09-950-022-1	Sequence 1, Appl
33	334.5	14.3	1873	13	US-09-950-022-3	Sequence 3, Appl
34	334.5	14.3	1873	13	US-09-950-022-5	Sequence 5, Appl
35	309	13.2	11527	14	US-10-108-605-70	Sequence 70, Appl
36	295.5	12.6	361	13	US-10-424-599-5459	Sequence 5459, Ap
37	231	9.9	411	13	US-10-424-599-44857	Sequence 44857, A
38	223.5	9.5	178	9	US-09-294-033B-4778	Sequence 4778, Ap
39	183	7.8	765	16	US-10-264-237-248	Sequence 248, App
40	159.5	6.8	373	13	US-10-424-599-55071	Sequence 55071, A
41	158	6.7	1446	10	US-09-822-846-559	Sequence 559, App
42	149	6.4	175	13	US-10-424-599-67266	Sequence 67266, A
43	133	5.7	917	13	US-10-424-599-44865	Sequence 44865, A
44	132	5.6	561	13	US-10-424-599-94913	Sequence 94913, A
45	130.5	5.6	602	15	US-10-106-698-2295	Sequence 2295, Ap

ALIGNMENTS

RESULT 1
US-10-425-114-1948
/ Sequence 1948, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jingdong

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Screen, Steven E

/ APPLICANT: Tabaska, Jack E

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(53313)B

/ CURRENT APPLICATION NUMBER: US/10425114

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 73128

/ SEQ ID NO 1948

/ LENGTH: 1844

/ TYPE: DNA

/ ORGANISM: Zea mays

/ FEATURE:

/ OTHER INFORMATION: Clone ID: 700172325_FLI

US-10-425-114-1948

Alignment Scores:

Pred. No.: 3,65e-268 Length: 1844
 Score: 2231.00 Matches: 429
 Percent Similarity: 97.33% Conservativity: 8
 Best Local Similarity: 95.55% Mismatches: 12
 Query Match: 95.26% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-1948 (1-1844)

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 Db 206 GAACATTGGCGATGCTCTCGTGAAGGTCGCCACTGTATTTCAGGCTATTTCGAGC 265
 QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
 Db 266 CTGCTCCAGGATTCATGAGTACAACTTCTATGTGACCGGGAGTGGCGTCATGATGAG 325
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 Db 326 CGCAACACCTACTATATCTGGGAGTTCGGTATAGTTAAACACACTATATTCGACAGGAA 385
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 Db 386 TTATATCAATAAATGCTTATTAAATCCAAAGCACACCTCGAAGCAGGATGAACATGGAT 445
 QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 Db 446 GTGGATAATGMAACTTTCAACATACGTTACATTTGCGATGTACCATTCAGAAAGGT 505
 QY 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 Db 506 ACTGTGAGAGTTTCAGAGGCTGCAATACAAATCTCAAGTTCGCGCTTCTGATATCTC 565
 QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
 Db 566 AATTTCATACATGCTATGATTTACTCCCGGATTCGGCAAGGTTATTGCTCTGACATT 625
 QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 Db 626 AATTTCATGAGCAATCTTTTCAATCTGCTGATGAACAGGGATTCCTGTAGTCTCT 685
 QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 Db 686 CTCTGGGACTATTTCAGAGGCTCAATTTGCGGCTTCGAGCCCATGGAATTCATCT 745
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
 Db 746 ATATTGCGGAGCTAGAAACTCATGGCTCGAACTTGACAGAGATCAGCTTGAACACAC 805
 QY 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
 Db 806 ACTATATCTGCATCGAAAGAGGCTAAGCGCAAACTTGTGAAGAAATGATGTCAGTGG 865
 QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
 Db 866 CGACACATCAGCATCTAGTGCATGCGCCACCCCTTATGAGTCTTCGAGGACATTCAGCTA 925
 QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGly 262
 Db 926 AAGCTTTTGCTAAATGACATTTTCACAGTCCAGTTATTTATTCATCATCATCAGATGGA 985
 QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
 Db 986 TCATTCCCTCAGTTATTGACCTTGCATCCCTTCTCGAATTTGAAATGATTTATTAGG 1045
 QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 Db 1046 TATTTTAAAAAATCAACTGTTAAATTTGCTTATTTCTGAACCAACCGGTGTCTCAATTCG 1195
 QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322

Db 1106 CTGGTTCTCTGGTTCCGAAATCGTGATCCAAACAGTGTGCTATGGCTATGTGGA 1165
 QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
 Db 1166 CTTAATGCATCATTAGTCTGCGCTTAACATGTGGTTCGAAGCTGGAGTGAGCTCAATA 1225
 QY 343 ProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 Db 1226 CCAATTGTGGATGAAACGACTCCCTGCTTGACACTTACTCTAGAAGTGCATCACAGCC 1285
 QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
 Db 1286 CTAGCTAAAGACAAAGGTCTACACATGTTCCGCTGGATGAGTACCAATTCATCAGCT 1345
 QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
 Db 1346 TTACAGCTTGACACAGATGCCATACGCCCTTTTGGATTTTTTAACGGTTCAGAGATGCCAG 1405
 QY 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
 Db 1406 ATGTGCCCTCCGCTCTGATCCTTTGCTGAAGGTGATGAGCGACTGGCTAATCCGGGGTG 1465
 QY 423 ArgArgValPheIleValGlnAlaGlySerLysArgValGluGlyIleIleSerLeuSer 442
 Db 1466 CGCGGGTGTTCATAGTAGAAGCTGGAGCAACAGTGTGAGGGTATTATATCACTGAGT 1525
 QY 443 AspIlePheLysPheLeuLeuSerLeu 451
 Db 1526 GATATTTTCAAGTCTTGTGAGCTTG 1552

RESULT 2

US-10-425-114-27250

; Sequence 27250, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 27250

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB4729-018-A12_FLI

US-10-425-114-27250

Alignment Scores:

Pred. No.: 6,69e-268 Length: 2228
 Score: 2230.00 Matches: 428
 Percent Similarity: 97.33% Conservativity: 9
 Best Local Similarity: 95.32% Mismatches: 12
 Query Match: 95.22% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-27250 (1-2228)

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 Db 589 GAACATTGGCGATGCTCTCGTGAAGGTCGCCACTGTATTTCAGGCTATTTCGAGC 648
 QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
 Db 649 CTGCTCCAGGATTCATGAGTACAGTTCATGTGTGACCGGGAGTGGCGTTCATGATGAG 708

QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 Db 709 CGCCAACTACTATATCTGGGAGTTCGGTATAGTTAAACACACTATATCTGACAGGGAA 768
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 Db 769 TTTATCAATTAATGAGCTTATTAATCAAGCAACCTGGAAGCAGGATGAACATGGAT 828
 QY 83 ValAspAsnGluAsnThrGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 Db 829 GTGGATAATGAAATTTCAACATACGTTACATGTCGATGATGATGATGATGATGATGAT 888
 QY 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 Db 889 ACTGTGAGAGTTTCAGAGGCTGCAATACAAATCTCAAGGTGCGCGTTCTGTAGTATCTC 948
 QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyValIleAlaLeuAspIle 142
 Db 949 AATTTCATACATGCTATGATTTACTCCGGATTCGGCAAGGTTATGCTCTGGACATT 1008
 QY 143 AsnLeuProValIysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 Db 1009 AATTACCTGTGAAGCAATCTTTCAATTCATGTCATGCAAGGGATTCCTGAGTCTCT 1068
 QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 Db 1069 CTCGTGGATCTCATTCAGAGTCAATTTGCGGCTTCGAGCCCATGGGATTTCAATCTT 1128
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
 Db 1129 ATATTGGGGAGCTAGAACTCATGGCTCGAATTCAGCAAGATCAGCTTGAACACAC 1188
 QY 203 ThrIleSerAlaTrpIysGluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
 Db 1189 ACTATATCTGCATGAAAGAGGCTAAGCGGCAAACTGTGGAAGAATGATGCTGAGTGG 1248
 QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
 Db 1249 CGAGCAGATCAGCATCTAGTGTGATGTCACCCCTTATGAGTCTTGAAGGACATTCAGTA 1308
 QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerGly 262
 Db 1309 AAGCTTTTGTAAATGACATTTCTACAGTCCAGGTTATTTATCATCATCATCATGATGSA 1368
 QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
 Db 1369 TCATTCCTCAGTATTGCACTTCATCCCTTCCTGGAATTTGAAATGATTTTATGG 1428
 QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 Db 1429 TATTTTAAATACTCACTGGTAAATTTGCTATTTCTGAACCAACCGGTGCTCCATTCGG 1488
 QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
 Db 1489 CTGGTTCTCTGGTTCCGAAATCGGTGATCCAAACAGTCGTCCATTCGGTATGTCGA 1548
 QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
 Db 1549 CCTAATGCATCACTTAGCTCTGCGCTTAAACATGTTGCTCAAGCTGGAGTGCATCATA 1608
 QY 343 ProfileValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 Db 1609 CCAATGAGTGAATGAACAGCTCCCTGCTGACACTTACTCTAGAAGTGACATCACAGCC 1668
 QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGluAla 382
 Db 1669 CTAGCTAAAGACAGGCTACACATGTTGCGCTGGATGAGATGACATTCATCAGGCT 1728
 QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
 Db 1729 TTACAGCTTGACAAAGATGCAATACGCTTTTGGATTTTAACTGATGATGATGATGATGAT 1788

QY 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
 Db 1789 ATGTCCTCCGGTCTGATCTCTTTCGTAAGGTGATGGAGCGACTGGCTAATCCGGGGTG 1848
 QY 423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer 442
 Db 1849 CGCGGGGTGTTCTGTAAGTAGAGCTGGGAGCAACGTTGGAGGGTATTATATCACTGAGT 1908
 QY 443 AspIlePheLysPheLeuLeuSerLeu 451
 Db 1909 GATATTTTCAAGTCTTCTGCTGAGCTTG 1935

RESULT 3

US-10-425-114-3817
 ; Sequence 3817, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 3817
 ; LENGTH: 1474
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700342215_FLI
 US-10-425-114-3817

Alignment Scores:

Pred. No.: 1,08e-261 Length: 1474
 Score: 2178.00 Matches: 421
 Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.53% Mismatches: 1
 Query Match: 93.00% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525c-2 (1-451) x US-10-425-114-3817 (1-1474)

QY 23 GluTyrLysPhePheValAspGlyGluTrpArgHisAspGluArgGlnProThrIleSer 48
 Db 2 GAGTACAAATCTTTGTGGACGGGGAATGGCGCATGATGAGCGTCAACCTACCATATCT 61
 QY 49 GlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyrAsnGlnIleAsnThr 68
 Db 62 GGGGAGTTTGGCATAGTTAAACACACTTTACTTGACAGGGAATATAACCAATTAACACC 121
 QY 69 LeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPhe 88
 Db 122 TTATCAGTCCAGCACACCTGGAGCAGGATGAACATGATGATGATGATGATGATGATGAT 181
 QY 89 GlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGlu 108
 Db 182 CAACGTACGGTTACGTTGTCAGATGGTACCGTTTCAGAGAGTACTCTGAGAGTTTCAGAG 241
 QY 109 AlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyr 128
 Db 242 GCTCAATACAAATATCTAGTGTGCTGTTCTCAATATCTGAGTTGTCATACATGCTAT 301
 QY 129 AspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProValLysGln 148
 Db 302 GATTTACTCCAGATCTGGCAAGGTTATTGCCCTAGACATTAATTTACCTGTGAGCAA 361
 QY 149 SerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTrpAspSerPheArg 168

```

Db      362 TCATCCATATCTCCATGAACAGGGATCTCTGTAGCTCCTCTCGGACTCATTGAGA 421
Qy      169 GlyGlnPheValGlyLeuLeuSerProLeuAspPheLeuLeuLeuLeuLeuLeuLeu 188
Db      422 GGTCATTTGTTGGTCTCTTAGCCCAATGGATTTCACTCATATTCGGGAGCTAGAA 481
Qy      189 ThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSerAlaTrpIys 208
Db      482 ACTCATGGCTCGAATCTTACAGAGAGACGCTTGAACACACACATATCTGCATGAAA 541
Qy      209 GluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLeu 228
Db      542 GAGCTAAGCGGCAAACTAATGAAGAAATGATGCTCAGTGGCGACCGCAACAGCATCTA 601
Qy      229 ValHisAlaThrProTyrGluSerLeuArgAspIleAlaValIysLeuLeuGlnAsnGly 248
Db      602 GTGCATGCCACCCCTTATGAGTCTTGAAGGACCTTGCAGTAAAGCTTTTGCAAAATGGC 661
Qy      249 IleSerThrValProValIleTyrSerSerSerSerAspGlySerPheProGlnLeuLeu 268
Db      662 ATTCTACAGTCCAGTTATTTATTCATCATCATCAGATGGATCATTCGGCGAGTTATG 721
Qy      269 HisLeuAlaSerLeuSerGlyIleLeuIysCysIleCysArgTyrPheIysAsnSerThr 288
Db      722 CATCTGTCATCACTTCTCGAATTTTGAATGTATTTGTAGATATCTCAAAACTCAACT 781
Qy      289 GlyAsnLeuProIleLeuAsnGlnProValCysSerIleProleuGlySerTrpValPro 308
Db      782 GGTAAATTTGCCATTCTGAACCAACAGTGTCTCTCATTCCTCGGTTCCTGGGTTCG 841
Qy      309 LysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSer 328
Db      842 AAAATGTGTGATCTGAACAGTGTCTCATTTGTTGCGACCTTAATGCATCATCTAGC 901
Qy      329 SerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAsn 348
Db      902 TCTGCCCTTACATGTTGGTTCAGCTGGAGTGAAGCTCAATACCAATTTGTGATGACAAC 961
Qy      349 AspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaIysAspIysVal 368
Db      962 GACTCCCTGCTTGACACTTACTTAGAAGTGAATCACAGCCCTTAGCAAAAGACAGGTC 1021
Qy      369 TyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGlnAsp 388
Db      1022 TACACATGTTTCGGCTGGAGTGAATGACCATTCACAGGCTTTGCAGCTTTGGACAAGAT 1081
Qy      389 AlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSerAsp 408
Db      1082 GCCAATACACTTTTGGATTTTAAACGGCCAGAGATGCCAGATGTGCTCCGCTCTGAT 1141
Qy      409 ProLeuLeuIysValMetGluArgLeuAlaAsnProGlyValArgArgValPheIleVal 428
Db      1142 CCTTCTGTAAGTGTATGGAGCGACTGGCTAATCTCGGTGCGGCGGTTCATTGTG 1201
Qy      429 GluAlaGlySerIysArgValGluGlyIleIleSerLeuSerAspIlePheIysPheLeu 448
Db      1202 GAAGCTGGAGGCAACGTGTGGAGGGGATCATATCATCACTAAGTGAATTTCAAGTCTTG 1261
Qy      449 LeuSerLeu 451
Db      1262 CTGAGCTTG 1270

```

RESULT 4

US-10-437-963-2873

; Sequence 2873, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

```

; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 2873
; LENGTH: 2980
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT4530_102608C.1
; US-10-437-963-2873

```

Alignment Scores:

```

Pred. No.: 6,99e-232 Length: 2980
Score: 1944.00 Matches: 361
Percent Similarity: 94.50% Conservative: 34
Best Local Similarity: 86.36% Mismatches: 23
Query Match: 83.01% Indels: 0
DB: 17 Gaps: 0

```

US-09-857-525C-2 (1-451) x US-10-437-963-2873 (1-2980)

```

Qy      3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db      685 GAACATTTGCCAATGTCTCTGTGAAGGTTGTCCACCGTTTTTCAGGCCATATCGAGT 744
Qy      23 LeuSerProGlyIleHisGluTyrIysPhePheValAspGlyGluTrpArgHisAspGlu 42
Db      745 TTGTCCCTCGGATTCATCAGTACAAATTTGTGTGGATGGGAATGGCGGCACGATGAG 804
Qy      43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrIleThrArgGlu 62
Db      805 CGGCACACTCATCAACAGGAGACTATGTTGTGGTGTGTTAACTTTATGCTTAGTAGGGAC 864
Qy      63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
Db      865 TTGACCCAGATTAATACATATTTGAGCCCTAGTACACCTGGAGTAGGATGAACATGGAT 924
Qy      83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
Db      925 GTGGACACAGCAATTTTCAACGTACTGTTTCTTGTCTGATGGCATTTATTCAGGAAGGT 984
Qy      103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrIleu 122
Db      985 CCTCAGGAATTTTCAGAGGCGAGCTATACAGATCTCTAGGTGTCTGTAGCGGATTTCGT 1044
Qy      123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyIysValIleAlaLeuAspIle 142
Db      1045 AATGGCAAACTGGTATGATTACTCCAGATCTTGGCAAGGTTCATGCTCTAGACGTT 1104
Qy      143 AsnLeuProValIysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db      1105 AATTTCCCTGTGAAGCAATCTTTTCATATTTCTTCATGATTCATGACAGGGAATTCCTGTGGACCT 1164
Qy      163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db      1165 CTATGGGATTCATTCAGGGGCCAGTTTGTGGCTTTTGGCCCTTGTAGCCCTGATTTTATCTC 1224
Qy      183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
Db      1225 ATATTGAGAGAGCTGCAAACTCATGGCTCCATCTGACAGAGAGAGCGCTTGAAACACAT 1284
Qy      203 ThrIleSerAlaTrpIysGluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
Db      1285 ACTATATCTGGGTGAAGAGGAGCGCAAGCGTCAAACTTATGCCAGAAACGAGGGTCTTGG 1344
Qy      223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
Db      1345 AGGGCAATTCACATTTAGTTTCATGCCACCCCTTATGAATCCCTGAGGGAATTCGTATG 1404

```

```

QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGly 262
DB 1405 AGATACGCAAAACGGTGTCTACCGTCCAAATATGTTTCTCATCCAGATGGC 1464
QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
DB 1465 TCATATCCCAAAATGTTGCATCTTGCTTCCCTTCGGAAATTTTGAATGTATTGTAGA 1524
QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
DB 1525 TATTTTAAATAATTTCTCAAGGTAAATTTACCTATTTGAGCCACCTGTATGCAATCTCT 1584
QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
DB 1585 CTGGGTACCTGGGTTCCTCAAAATTTGGTATCTTAATGGCGTCCATTTGGCTATGTTGGG 1644
QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerIle 342
DB 1645 CCTAACACATCTCTTAGCGCTGCCCTGAAATTTGCTGGTTCAAGCTGGTGTGAGCTCAATA 1704
QY 343 ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
DB 1705 CCAATTTGGATGACAAATGACTCACTGCTTGACACATATTCAGAAAGTGCATCACAGCT 1764
QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
DB 1765 CTTGCANAAACAAAGGTGTACACACATTCGCTTAGATGAGATGACCAATTCATCAGGCT 1824
QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
DB 1825 CTGAGCTTGCACAGGATGCAAAATTCGCCCTTTGGATTCCTTAACGGCAAGATGTCNG 1884
QY 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnPro 420
DB 1885 ATGTGTCTCGGCTGCACACTTTGTTGAAAGTGATGAGCGGATTAGCTAATCTCT 1938

RESULT 5
US-10-425-114-33827
: Sequence 33827, Application US/10425114
: Publication No. US20040034888A1
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E.
: APPLICANT: Tabaska, Jack E.
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(5313)B
: CURRENT APPLICATION NUMBER: US/10/425,114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 33827
: LENGTH: 1387
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURES:
: OTHER INFORMATION: Clone ID: UC-ZMFLM017174G09_FLI
US-10-425-114-33827

Alignment Scores:
Pred. No.: 1.16e-214 Length: 1387
Score: 1802.00 Matches: 352
Percent Similarity: 97.53% Conservative: 4
Best Local Similarity: 96.44% Mismatches: 9
Query Match: 76.94% Indels: 0
DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-33827 (1-3387)
QY 87 AsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgVal 106

```

```

DB 3 AATTTTCAACATACGGTTACATTGTGGATGGTACCAATCCAGAAAGTACTGTGAGAGTT 62
QY 107 SerGluAlaAlaIleGlnIleSerArgCysArgValSerGluCysIleLeuAsnLeuHisThr 126
DB 63 TCAGAGGCTGCAATACAAATCTCAAGGTCGCGGTTTCTGAGTATCTCAATTTGCAATCA 122
QY 127 CysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProVal 146
DB 123 TGTATGATGTTTATCTCCGGAATTCGGCAAGGTATTATGCTCTGCACATTAATTTACCTGTG 182
QY 147 LysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaProLeuTrpAspSer 166
DB 183 AAGCAATCTTTTCAATTTCTGCATGAACAGGGGATTCCTGTAGTCTCTCTGGACTCA 242
QY 167 PheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGlu 186
DB 243 TTCAGAGGTCAATTTGTGGCCCTCTGAGCCCATTTGGAATTTCACTTATTTGCGGGAG 302
QY 187 LeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerAla 206
DB 303 CTGAGAACTCTGATGGCTCGAATCTGACAGAGATCAGCTTGAACACACACATATATCTGCA 362
QY 207 TrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGln 226
DB 363 TGGAAAGAGGCTAAGCGGCAACTTGTGGAAGAAATGATGTCAGTGGCGAGCACATCAG 422
QY 227 HisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeuGln 246
DB 423 CATCTAGTGCATGCCACCCCTTATGAGTCTCTTGGAGGACATTCAGTAAGCTTTTGGCAA 482
QY 247 AsnGlyIleSerThrValProValIleTyrSerSerSerAspGlySerPheProGln 266
DB 483 ANTGCATTTCTACAGTGCACCTTATTTATTCATCATCATCAGATGATCCTCCTCAG 542
QY 267 LeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyr2helysAsn 286
DB 543 TTATTTGACCTTGCATCCCTTTCTGGAAATTTTGAATGTATTTTAGGTATTTTAAAAAC 602
QY 287 SerThrGlyAsnLeuProIleLeuAspGlnProValCysSerIleProLeuGlySerTrp 306
DB 603 TCAACTGGTAAATTTGCCATTTCTGAACCAACCGGTGTGCTCCATTCGCTGGGTCTCGG 562
QY 307 ValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSer 326
DB 663 GTTCCGAAATTCGGTATCCAAACAGTCGTCATTTGGCTATGTTGCGACCTAATGCAATCA 722
QY 327 LeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAsp 346
DB 723 CTTAGCTCTGCCCTTAAATGTTGGTTCAAGCTGGAGTGAGCTCAATACCAATTTGGAT 782
QY 347 AspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaLysAsp 366
DB 783 GAAACGACTCCCTGCTTGCACATTTACTCTAGAGTGACATCACACCCCTAGCTAAGAC 842
QY 367 LysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGly 386
DB 843 AAGGTCTACACACATGTTGCTGCTGGATGAGATGACCAATTCATCAGGCTTTTACAGCTTGA 902
QY 387 GlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArg 406
DB 903 CAAGATGCCAATACGCTTTTGGATTTTAAACGGTTCAGAGATGCCAGATGTCCTCGG 962
QY 407 SerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgArgValPhe 426
DB 963 TCTGATCCTTTGCTGAAGGTGATGAGCGACTGCTGCTATCCGGGGGTGCGCGGGGTGTC 1022
QY 427 IleValGluAlaGlySerLysArgValGluGlyIleLeuSerLeuSerAspIlePheLys 446
DB 1023 ATAGTAGAAGCTGGAGCAACGCTGTGAGGGGTATATATCATCTGAGTGATATTTTCAAG 1082
QY 447 PheLeuLeuSerLeu 451

```


Db 1083 TTCTTGCTGAGCTTG 1097

RESULT 6

US-10-424-599-44858

; Sequence 44858, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 44858

; LENGTH: 3077

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT MRT3847_140506C.1

US-10-424-599-44858

Alignment Scores:

Pred. No.:	1-58e-181	Length:	3077
Score:	1542.00	Matches:	306
Percent Similarity:	80.62%	Conservative:	60
Best Local Similarity:	67.40%	Mismatches:	75
Query Match:	65.84%	Indels:	13
DB:	13	Gaps:	4

US-09-857-525C-2 (1-451) x US-10-424-599-44858 (1-3077)

QY 3 GluHisLeuProMetSerProleGlyCysProThrValPheGlnAlaLeuCysSer 22

Db 1019 GAACCTTCTACAAATGTCGCCAGTGAAGGTTGTCCAACTGTTTCAAGTGAATCATAGC 1078

QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluThrArgHisAspGlu 42

Db 1079 TTGCTACCGGTTACCACTACAGTCAAGTTTGTGTGAGAGATGCGGCATGATGA 1138

QY 43 ArgGlnProThrIleSer-GlyGluPheGlyIleVal-AsnThrLeu-TyrLeuThrArg 61

Db 1139 CATCAACTTATGAGCTGAGATATGGATAGTCTAACACGTCTTATTTGGCCACT 1198

QY 62 GluTyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMet 81

Db 1199 GATCCTAAATACATACCTGTTTAA-----CTCCAGACGTTGCTTCTGGAATAGCATG 1252

QY 82 AspValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGlu 101

Db 1253 GATGTGATATGATGCTTTTCCCGAATGCGCGGTGACCGATGCTTGTGATGAGTGAG 1312

QY 102 GlyThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyr 121

Db 1313 GTGCTGCCAGATATACAGATCTGATGATCAAAATATCCCGTCAAGTATTTCTGCATTT 1372

QY 122 LeuAsnLeuHisThrCysTyrAspLeuProAspSerGlyLysValIleAlaLeuAsp 141

Db 1373 CTATCTTCACACCCGCTTATGAATTAATCTCCGAGTCAAGGATGTTGTGTTGGAT 1432

QY 142 IleAsnLeuProValIysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAla 151

Db 1433 GTTGATCTACCACTGAACACGCAATTTATATTCATATTCATGAGCAGGATTTTCATGGCT 1492

QY 162 ProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIle 181

Db 1493 CCTCTTTGGGACTCTCTGCAAGGGCAATTTGTTGGTGTCTAGTCTTCGGATTTAT 1552

QY 182 LeuIleLeuArgGluLeuThrHisGlySerAsnLeuThrGluGluGlnLeuGluThr 201

Db 1553 TTAATTTTAAAGAGAGCTGGGAATCATGATCCCAATCTGACAGAGAGGAGCTTGAACA 1612

QY 202 HisThrIleSerAlaTrpLysGlyAlaLys-----ArgGlnThrAsnGlyArg 217

Db 1613 CATACCATATCAGCTTGAAGAGAGGAAATCGTATCTAAATATGACAGAACATGGACAT 1672

QY 218 AsnAspSerGlnTrpArgProGlnHisIleValHisLeuValHisAlaTrpProTyrGluSerLeu 237

Db 1673 GGAACCTGCATTTCAAGA-----TGTTTATCCATCAGGCGCATATGATTAATCTG 1723

QY 238 ArgAspIleAlaValIysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSer 257

Db 1724 AAGATATTTGCCATGAAGATCTTGCAAAAGGAGGTTTCAACAGTCTCTATATTCATCA 1783

QY 258 SerSerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeu 277

Db 1784 TCTTCTGAAGATGCTTCATTTCCACAGTTACTACATCTTGCTTCACTTTCAGGAATATT 1843

QY 278 LysCysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnPro 297

Db 1844 AAATGCATTTGTAGTATTTTAGGCACTCTCTAGTTCCTTGCTCTTCACTTCACTTCCA 1903

QY 298 ValCysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgPro 317

Db 1904 ATCTGTCAATACCTGTGGGCACTGGGTGCCAAATTTGGGAATCAANTCGCGGCT 1963

QY 318 LeuAlaMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAla 337

Db 1964 CTAGCAATGTTGAGACCAACCGCTTCTCTGTGCGTCAAGCTTAAATTTATTAGTTCAAGCC 2023

QY 338 GlyValSerSerIleProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArg 357

Db 2024 CAAAGTAAGCTCAATACCAATAGTTCATGATATGACTCATTTATGATATATATCTGCTCG 2083

QY 358 SerAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMet 377

Db 2084 AGTGACATACAGCTTTTGGCAAGAACAGAGCATATACATATTAATCTTCGCAAGATG 2143

QY 378 ThrIleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsn 397

Db 2144 ACTGTTTCATCAGCTTTGAGTTGCGTGGCCAGACCATATAGTCCCTATAGAG---CTTAGA 2200

QY 398 GlyGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeu 417

Db 2201 AGTCAAGATGTCAGATGTTTGGGTTCTGATCTCTGCAATTAAGTATGGAACGCTTG 2260

QY 418 AlaAsnProGlyValArgValPheIleValGluAlaGlySerLysArgValGluGly 437

Db 2261 GCAATCCAGTGTCAAGCGCTTGTGATCGTGAAGCTGGCAGCAAGCGGTGAGAGGCG 2320

QY 438 IleIleSerLeuSerAspIlePheLysPheLeuLeu 449

Db 2321 ATGTTTTCATGAGTGCATATTCAGTTCTTCATT 2356

RESULT 7

US-10-425-114-6702

; Sequence 6702, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 6702

; LENGTH: 1200

; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700575557_FLI
 ; US-10-425-114-6702

Alignment Scores:

Pred. No.: 5,89e-176 Length: 1200
 Score: 1492.00 Matches: 292
 Percent Similarity: 97.35% Conservatives: 2
 Best Local Similarity: 96.69% Mismatches: 8
 Query Match: 63.71% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-6702 (1-1200)

QY 150 PheHisIleLeuHisGluGlnGlyProValAlaProLeuTrpAspSerPheArgGly 169
 DB 2 TTTCATATTCTGATGAACAGGGGATTCCTGTAGCTCCTCTCGGGACTCAITCAGAGGT 61
 QY 170 GlnPheValGlyLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGluThr 189
 DB 62 CAATTGTGCGGCTCTGAGCCATTGGATTTCATCTATATATGCGGAGCTAGAACT 121
 QY 190 HsGlySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerIlePheGlu 209
 DB 122 CATGGCTCGAATCTGACAGAGATGCTTGAACACACACTATATCTGCATGGAAAGAG 181
 QY 210 AlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLeuVal 229
 DB 182 GCTAAGCGCAACTGTGGAGAAATGATGTGTCAGTGGGAGCATCAGCATCTAGTG 241
 QY 230 HisAlaThrProTyGluSerLeuArgAspIleAlaValLysLeuGlnAsnGlyIle 249
 DB 242 CATGCCACCCCTTATGAGTCTCTGAGGACATTCGAGTAAAGCTTTGCTAANTGACAT 301
 QY 250 SerThrValProValIleTySerSerSerSerAspGlySerPheProGlnLeuHis 269
 DB 302 TCTACAGTGCAGATTATTATTCATCATCATGATGATCATCCCTCAGTTATTGAC 361
 QY 270 LeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyPheLysAsnSerThrGly 289
 DB 362 CTTCATCCCTTTCTGGAATTTGAATGTATTTTAGTATTTTAAAACTCAACTGT 421
 QY 290 AsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpValProLys 309
 DB 422 AATTGTGCTATTCTGAACCAACCGGTGTGCTCCATTCGGTTCCTGGTTCGGAA 481
 QY 310 IleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerSer 329
 DB 482 ATCGGTGATCCAAACAGTCTGCTTCCTATGCTATGTCGACCTAATGCACTAGCTCT 541
 QY 330 AlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAspAsnAsp 349
 DB 542 SCCCTTAACATGTGTGTTCAAGCTGGAGTGAGTCAATACCAATTTGGATGAAACGAC 601
 QY 350 SerLeuLeuAspThrTySerArgSerAspIleThrAlaLeuAlaLysAspLysValTy 369
 DB 602 TCCCTGTCTGACACTTACTCTAGAAGTGACATCACAGCCCTAGCTAAAGACAAAGTCTAC 661
 QY 370 ThrHisValArgLeuAspGluMetThrIleHsGlnAlaLeuGlnLeuGlyGlnAspAla 389
 DB 662 ACATATGTTGCTTGATGATGATGACCATTCATCAGGCTTTACAGCTTGGACAAAGATGCC 721
 QY 390 AsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSerAspPro 409
 DB 722 AATACGCTTTTGGATTTTAAACGTCACAGATGCCAGATGTCCCTCCCGTCTGATCCT 781
 QY 410 LeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgValPheIleValGlu 429
 DB 782 TTCTGAAGGTGATGAGCGACTGGGTAAATCCGGGGGTGGGGGGTTCATAGTAGAA 841
 QY 430 AlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPheLeuLeu 449

Db 842 CTGCGAGCAACGTTGGAGGTATTATATCACTAGTGATATTTCAGTTCTTGCTG 901
 QY 450 SerLeu 451
 Db 902 AGCTTG 907
 RESULT 8
 US-10-424-599-79689
 ; Sequence 79689, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 79689
 ; LENGTH: 2315
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_42975C.1
 ; US-10-424-599-79689

Alignment Scores:

Pred. No.: 7.51e-175 Length: 2315
 Score: 1487.00 Matches: 287
 Percent Similarity: 78.05% Conservatives: 65
 Best Local Similarity: 63.64% Mismatches: 75
 Query Match: 63.49% Indels: 24
 DB: 13 Gaps: 4

US-09-857-525C-2 (1-451) x US-10-424-599-79689 (1-2315)

QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
 DB 585 CATATAGTATGCTCTCCGATGCGGATGCGCTGCTGTTTCAAGTTGTTGCAACTTA 644
 QY 24 SerProGlyIleHisGluTyLysPhePheValAspGlyGluTrpArgHisAspGluArg 43
 DB 645 ATGCCGGGGTTTCATCAGTACAAATTAATGTAGATGTGTAGTGGCGGCACCATGASCAG 704
 QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyLeuThrArgGluTy 63
 DB 705 CAGCATTTGTAATGGAGGAGCTGTGGAGTAGTGAACTATTTATATATAGTAGAGAACCA 764
 QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
 DB 765 GATATCTTACCTCCATCTTAATAACTGAAACACTGCT---CGATCACATGAGGTT 821
 QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
 DB 822 GACCACTGGAAGCTAATCCAGCATGACTGTGCTCAT----- 860
 QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyLeuAsn 123
 DB 861 -----CTGGAGTTTCTCGCCCATATATATGCTTCTGTC 899
 QY 124 LeuHisThrCysTyAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
 DB 900 ACGCATATCTCATATGATTTCCTCCGAGTCCAGAAAGGTCAITGCTTGGATATATAAT 959
 QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
 DB 960 TTACCAAGTTAAGCAAGCATTCCTCATGTTCTTTACGAAACAGGCTATATCTATGGCTCCTCTA 1019
 QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuLeu 183

```

1020 TGGGATTTTTCAGGAGCAGTTCCTGGAGTCTTAGTGCATGAGCTCATTCURATA 1079
184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThr 203
1080 TTGAAGAGCTGGGACCATAGTTCATATTTGACCTCAAGACCACTCGAGACTCATACT 1139
204 IleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln----- 221
1140 ATAGTCTGCTGGAAGAGCAATTTTCAACAACTTCAAGACACTTGCATAGTAATGGGGA 1199
222 -----TrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArg 238
1200 TCATATCTCTTGGCGA-----TTTGTTTCATGCGGACCCCATGAATGCTTAATA 1247
239 AspLeuAlaValLysLeuGlnHisGlySerThrValProValIleTyrSerSer 258
1248 GATGCGGCTTTGAGGTTTTCAGAAACAGGTGTCACACTGTCCTTATCATATCTTCT 1307
259 SerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLys 278
1308 TCAGAGGATGTTTCATATCTCTCACTGCTACATCTTGCTTCCCTATCCGGAATATAA 1367
279 CysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProVal 298
1368 GGTATATGCGGCACTTTTAAGCACTCTTGAAGTCTTGGCCATCTTCAACTTCCAGTG 1427
299 CysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeu 318
1428 GCTTCAATACCTTTGGGTACATGATGCTTAGAGTTGGGGAACAAATGTCGGCCACTA 1487
319 AlaMetLeuArgProAsnAlaSerLeuSerAlaLeuAsnMetLeuValGlnAlaGly 338
1488 GCATGTTGATGCCAAGTCTCTCTCGTGTCTGCTTGTCTATGTTTGTTCAGCCGAA 1547
339 ValSerSerIleProIleValAspAsnAspSerLeuLeuAsnThrTyrSerArgSer 358
1548 GTTAGCTCAATACCGATGTCGATGCTTAATGATGCTTGTGCTTGCATTTATTCAGAGC 1607
359 AspIleTrpAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThr 378
1508 GATATTAAGTGGCTTAAAGATAAAGCTTATGCTAGGATATCTCTGATGAAATAGT 1667
379 IleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGly 398
1668 ATTCACGAGCAATGCTTTGGGACAGATGCAACTTCTCTTAAGGCTTTACATGGT 1727
399 GlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAla 418
1728 CACAGATGTCACATGCTTTGAGATCTGATTCGCTTTTACAAAGTGATGGAGCGTTCGCT 1787
419 AsnProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIle 438
1788 AATCCTGGGTAGGACATTTGGCTGTGGAGCTGCGCAGCAAGCGTGTGGAAGGGATT 1847
439 IleSerLeuSerAspIlePheLysPheLeuLeu 449
1848 ATTTCTTTAAGCATGGTTTCAGATCTTGTGTA 1880

```

RESULT 9

```

US-10-425-114-29527
; Sequence 29527, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B

```

```

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29527
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY057G02_FLI
US-10-425-114-29527

```

```

Alignment Scores:
Pred. No.: 6,99e-155 Length: 1507
Score: 1325.00 Matches: 256
Percent Similarity: 83.24% Conservative: 52
Best Local Similarity: 69.19% Mismatches: 54
Query Match: 56.58% Indels: 8
DB: 13 Gaps: 3

```

US-09-857-525C-2 (1-451) x US-10-425-114-29527 (1-1507)

```

QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
DB 3 GATAATGAGCTTTTCGCCGAATGCCCGGTGACGATGGTACTTTGAGTGGGTCGTG 62
QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsn 123
DB 63 CCAAGATATCAGATCTGATGATCAATATCCCGTCAGCGTATTTCTGCAATCTCTCT 122
QY 124 LeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
DB 123 TCACACACCGCTTATGAATTAATCTCCGAGTCAGCAAGGTGTGCTTTGATGTTGAT 182
QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
DB 183 CTACAGCTGAAACAGGCATTTTCATATATTCATATGATGAGCAGGAGGTTTTCATGCTCTCT 242
QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
DB 243 TGGGACTTCTGCAAGGGGCAATTTGTTGTTCTTCTAGTCTGCTGATTTTATTTAAT 302
QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThr 203
DB 303 TTAAGAGAGCTGGGGAATCATGGAATCCAAATCTGACAGAGAGGAGCTTTGAACACATACC 362
QY 204 IleSerAlaTrpLysGluAlaLys-----ArgGlnThrAspGlyArgAsnAsp 219
DB 363 ATATCAGCTTGGAAAGAGGAATTCGTATCTAAATAGACAGACAAATGGACATGGAACT 422
QY 220 SerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAsp 239
DB 423 GCATTTTCAAGA-----TGTTTATCCATGAGGCGCATATGATAATCTGAAAGAT 473
QY 240 IleAlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSer 259
DB 474 ATTGCCATGAGATCTTGGAAAGAGGTTTCACTGTTCTTATTCATTCATCTCTCT 533
QY 260 SerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCys 279
DB 534 GAAGATGCTTCAATTTCCACAGTTACTACATCTTGCTTCACTTTCAGGAATACTTAATGC 593
QY 280 IleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCys 299
DB 594 ATTGTAGGTATTTTAGGCACTCTCTAGTTCCTTGTCTGCTTCACTTCAACTTCCAATCTGT 653
QY 300 SerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAla 319
DB 654 GCAATACCTCTGGGACAGTGGGTGCCCAAAATTTGGGAATCAAAATCGCGGCTCTTAGCA 713
QY 320 MetLeuArgProAsnAlaSerLeuSerAlaLeuAsnMetLeuValGlnAlaGlyVal 339
DB 714 ATGTTGAGACCAACCGCTTCTCTTGGTTCAGCCCTAAATTTATTAGTTCAAGCCCAAGTA 773

```

QY 340 SerSerLeuProIleValAspAspSerLeuLeuAspThrTyrSerArgSerAsp 359
 DB 774 AGCTCAATACCAATAGTATGATGATAATGACTCATTTATTTGGATATATATCTGCGAGTGAC 833
 QY 360 IleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIle 379
 DB 834 ATAAACAGCTTTGGCAAGAACAGACGATATACATATTAATCTTTGACGAATGACTGTT 893
 QY 380 HisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGln 399
 DB 894 CATCAGGCTTTGCGAGTTGGCGCAGACGATATAGTCCCTATGAG---CTTAGAAGTCAA 950
 QY 400 ArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsn 419
 DB 951 AGATGTCAGATGTCGCTGCTGATCCTCTGCACTAAAGTGATGGAACGCTTGGCAAT 1010
 QY 420 ProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIle 439
 DB 1011 CCAGGTGTCAGCGCGCTTGTCATCGTGAAGCTGCGCAGCAAGCGTGTAGAAGGCATTGTT 1070
 QY 440 SerLeuSerAspIlePheLysPheLeuLeu 449
 DB 1071 TCATTGAGTGACATATTCAAGTTCTTCATT 1100

RESULT 10

US-10-424-599-105
 ; Sequence 105, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 105
 ; LENGTH: 1879
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_100095C.1
 US-10-424-599-105

Alignment Scores:
 Pred. No.: 1,4e-104 Length: 1879
 Score: 924.50 Matches: 196
 Percent Similarity: 65.35% Conservative: 36
 Best Local Similarity: 55.21% Mismatches: 56
 Query Match: 39.47% Indels: 67
 DB: 13 Gaps: 6

US-09-857-525C-2 [1-451] x US-10-424-599-105 (1-1879)

QY 128 TyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProValLys 147
 DB 447 TATCGACATTTCCAGATGCGTATAAAGTGCACATTCGATTCTGAATCACTCATCTATCTT 506
 QY 148 GlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTrpAspSerPhe 167
 DB 507 CAATATTGGTAATGCTTGTAAACTTGTGTCTT----- 542
 QY 168 ArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGluLeu 187
 DB 543 -----TTTCTTTGAAACACAGGCTC 563
 QY 188 GluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerAlaTrp 207
 DB 564 GGAAATCATGGTGCTCAATCTAACAGAGAGGAGCTTGAACACATACCATATCAGCTTGG 623

QY 208 LysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgPro---GlnGln 226
 DB 624 AAAGGAGGAAAA-----TGGACAGATTTCACAA 653
 QY 227 HisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuGln 246
 DB 654 TGTTTATCCGTCAGGCGCATACGATAATTTGAAAGAGATTGCTGTGAAGATCCTGCAA 713
 QY 247 AsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGlySerPheProGln 266
 DB 714 CATGAAATTTCAACAGTTCTCTATTATTCAT-----TCAGAGATGGTTCAITTCACAG 767
 QY 267 LeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsn 286
 DB 768 CTACTACATCTTGCTTCACTTTCAGAAATCCTTAATGCAATTCAGGGTATTTTAGGAA 827
 QY 287 SerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrp 306
 DB 828 TGTTCATGTTTCATTCGCTTATACCTTCAACTTCCAAATTTGTGCAATCCCTGTGGCAGTGG 887
 QY 307 ValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSer 326
 DB 888 GTGCCCAAAATTTGGGAAATCAAAATGGCAGGCTCTAGCAATGTTGAGACCAACCGCTTCT 947
 QY 327 LeuSerSerAlaLeuAsnMetLeuValGln----- 336
 DB 948 CTGCGTCAGCCCTAAATTTATTTAGTTCAGGTGTGAACCTTTTGTCTGACCTTGAACAT 1007
 QY 336 ----- 336
 DB 1008 AGTGTATCAAAAAGAAAGCAACAGAGAGAAAGTGAAGAAATTTGGTATCCTATTATAC 1067
 QY 337 -----AlaGlyValSerSerIleProIleValAspAspAspSerLeuLeuAspThr 354
 DB 1068 TTCTGAGCCCAAGTAAGCTCAATACCAATAGTTCATGATATATGATCAATTAATCTACTGATATA 1127
 QY 355 TyrSerArgSerAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeu 374
 DB 1128 TATTGTGCGAGTGACATACAGCTTTGGCAAGAACAGAGCATATGCAATTAATCTT 1187
 QY 375 AspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGly 394
 DB 1188 GATGAATGACAGTTTCATCAGGCTTTCAGTTGGCGCAGGACGATATAGTCCCTATGAG 1247
 QY 395 PhePheAsnGlyGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMet 414
 DB 1248 ---CTTAGAAGTCAAGATGTCAGATGTGTTTGGTCTCTGATCCTCTGCAATAAAGTGTATG 1304
 QY 415 GluArgLeuAlaAsnProGlyValArgArgValPheIleValGluAlaGlySerLysArg 434
 DB 1305 GAACGCTTGGCAAAATCCAGGTGTCAGGCGCTTGTCTCATCGTGAAGCTGCGCAAGCGT 1364
 QY 435 ValGluGlyIleIleSerLeuSerAspIlePheLysPheLeuLeu 449
 DB 1365 GTAGAAGGCATTGTTTCAITTCAGGAGACATATTCAAGTTCTTCATT 1409

RESULT 11

US-10-424-599-101
 ; Sequence 101, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 101
 ; LENGTH: 1225

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(1225)

OTHER INFORMATION: unsure at all n locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_100091C.1

US-10-424-599-101

Alignment Scores:

Pred. No.: 1.89e-104 Length: 1225

Score: 521.00 Matches: 186

Percent Similarity: 81.50% Conservative: 31

Best Local Similarity: 69.92% Mismatches: 34

Query Match: 39.33% Indels: 15

DB: 13 Gaps: 4

US-09-857-525C-2 (1-451) x US-10-424-599-101 (1-1225)

QY 186 GluLeuGluThrHisGlySerAsnLeuThrGluGluGluLeuLeuGluThrHisThrIleSer 205
 Db 21 CAGCTTGAATCATGGTCCCAATCTACAGAGAGGAGCTTGAACACATACCATCA 80
 QY 206 AlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgPro--- 224
 Db 81 GCTTGGAAAGAGGAGGAAA-----TGACAGGATTT 110
 QY 225 GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeu 244
 Db 111 ACACAATGTTTATCGGTCAGGGCCATACGATAATTTGAAGAGATTGCTGTGAAGATC 170
 QY 245 LeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGlySerPhe 264
 Db 171 CTGCAACATGGAATTTCAACATGCTCTATTATCAT-----TCGAAGATGGTTCATT 224
 QY 265 ProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPhe 284
 Db 225 CCACAGCTACTACATCTTGCTTCACTTCAGGAATCCTTAATGATCATTTGAGGTATTT 284
 QY 285 LysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGly 304
 Db 285 AGGAATGTTTCTAGTTCATGCTTACTTCAACTTCAACTTTCGAATTTGCAATCCCTGTGGC 344
 QY 305 SerTrpValProLysIleGlyLeuAsnSerArgProLeuAlaMetLeuArgProAsn 324
 Db 345 ACGTGGGTGCCCAAAATTTGGGAGTCAAAATCGCGGCTCTAGCAATGTTGAGACCAAT 404
 QY 325 AlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIle 344
 Db 405 GCTTCACTTACTTCAGCCCTAAACTATTAGTTCAAGCTCAAGTAAAGTTCAATACCAATA 464
 QY 345 ValAspAsnAspSerLeuLeuAsp-ThrTyrSerArgSerAspIleThrAlaLeuAl 364
 Db 465 GTTGATGATAGTACTTACTTACTGATATTTATTTGTCGAAGTATATAACAGCTTTGGC 524
 QY 364 aLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuG1 384
 Db 525 AAAGACAGAACTTATACCCATATTAACTTGTGATGAATGACTGTTCAATCAGGCATTGA 584
 QY 384 nLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGlnMetCy 404
 Db 585 ATTGGGCAGGATTTCTTATATACTATTAG---CTGAGTTGTCAAGATGTCAGATGTG 641
 QY 404 sLeuArgSerAspProLeuLysValMetGluArgLeuAlaAsnProGlyValArgAr 424
 Db 642 TTTGCGAACTGATTCTCTGCAATAAGTGAAGAGGTTTGGCAATCCAGGTGTGAGCG 701
 QY 424 gValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIl 444
 Db 702 GCTTGAATTTGGAGCTGCGACAGCGGTAGAGGCATCATGCAATTTGAGTGACAT 761
 QY 444 ePheLysPheLeuLeu 449

Db 762 ATCAACTTCTTCCTT 777

RESULT 12

US-10-437-963-90730

Sequence 90730, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5321)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 90730

LENGTH: 6747

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_89371C.1

US-10-437-963-90730

Alignment Scores:

Pred. No.: 1.21e-101 Length: 6747

Score: 908.50 Matches: 203

Percent Similarity: 64.95% Conservative: 36

Best Local Similarity: 55.16% Mismatches: 92

Query Match: 38.79% Indels: 37

DB: 17 Gaps: 9

US-09-857-525C-2 (1-451) x US-10-437-963-90730 (1-6747)

QY 42 GluArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArg 61
 Db 4912 GAACCGGTACACGCAATTTCCAAACGCTACGGTTGGTAGAGGAGCCCTCTATCGCGC 4971
 QY 62 GluTyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMet 81
 Db 4972 GCGGTACACGCAATCTCTGAAGTGCATTCCTCAGGAACAAGCGCTTCTTGTCT 5031
 QY 82 AspValAspAsnGluAsn-----PheGlnArgThrValThrLeuSerAspGlyThr 98
 Db 5032 GACATCCATGAGGCGGAGTGGGAGCCCATTCGGCTCGCGCACCTGGTTGGCAGGCC 5091
 QY 99 ValSerGluGly-----ThrLeuArgValSerGluAlaIleGlnIleSerArg 115
 Db 5092 TTTGCTCAAGGATTTTATTGGCCGACAGCTCTCAATGATGACATCGATCTGTCTCGCGCA 5151
 QY 116 CysArgValSerGluTyr-----LeuAsnLeuHis-----ThrCysTyrAspLeu 130
 Db 5152 TGCAGAGCGTGTAGTCTCCAGCCAGCAATTCATCAGCCGCCGCCAGCCCTGCAGATC 5211
 QY 131 LeuProAspSerGly-----LysValIleAlaLeuAspIleAsnLeuProValLysGln 148
 Db 5212 ATACCACTGTATGCGCCATTTGCTGTCTGGGGGCTTGATATCTCTGGGACCGTTCAAACGG 5271
 QY 149 Ser-----PheHisIleLeuHis-----GluGln 156
 Db 5272 GCCCGCGCGGGTTTGAGTATCTGTATGTTGCGGTCCGACAAAGTTCAATGAGGCCGAG 5331
 QY 157 GlyIleProValAlaProLeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSer 176
 Db 5332 GCTTATCCGGTTGTAAG-----ATCGATAAGCACTCT 5364
 QY 177 ProLeuAspPheIleLeuLeuArgGlu-----LeuGluThrHisGlySerAsn 193

```

Db      5365  GCTCTAAATTCATTAAGGCGATCACGCGCGTTTGGACTGGAAATCTATGGCTCCAAT 5424
Qy      194  LeuThrGluGlnLeuGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGln 213
Db      5425  CTGACAGAAGAGCAGCTTGAACACATACCTATATCTCGGTGGAAGAGCGGCTCAA 5484
Qy      214  ThrAsnGlyArgAsnAspSerGlnTrpArgProGlnHisLeuValHisAlaLeuPro 233
Db      5485  ACTATGCCGAACAGAGGGTCTTGGAGGGCAATACCAATTTAGTTCTGCGCACCCCT 5544
Qy      234  TyrGluSerLeuArgAspIleAlaValLysLeuGlnAenGlyIleSerThrValPro 253
Db      5545  TATGAATCCCTGAGCGAAATGCTATGAGATACTGCAAAACGGTGTCTACCGTTCCA 5604
Qy      254  ValIleTyrSerSerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeu 273
Db      5605  ATTATGTTTTCCTCATCACAGATGGCTCATATCCCAATTTGTCATCTTGTCCCTT 5664
Qy      274  SerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIle 293
Db      5665  TCTGGAATTTGAATGTATTTAGATAATTTAAATAATCTCAAGTAAATTTACCTATT 5724
Qy      294  LeuAsnGlnProValCysSerIleProLeuGlySerTrpValProLysIleGlyAspLeu 313
Db      5725  TTGAGCCAACTGTATGACAAATTCCTCTGGGTACCTGGGTCCAAAATTTGGTGATCCT 5784
Qy      314  AsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerAlaLeuAsnMet 333
Db      5785  AATGGCGCTCATGGCTATGCTGGCGCTTAACACATCTCTTAGCGTGGCCCTCAATTTG 5844
Qy      334  LeuValGlnAlaGlyValSerIleProIleValAspAspAsnAspSerLeuLeuAsp 353
Db      5845  CTGTTCAAGCTGTTGAGCTCAATACCAATTTGGTGTGATGATGATGATGATGATGAT 5904
Qy      354  ThrTyrSerArgSerAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArg 373
Db      5905  ACATATTCAGAAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5964
Qy      374  LeuAspGluMetThrIleHisGln 381
Db      5965  CTAGATGAGATGACCATTCATCAG 5988

```

```

RESULT 13
US-10-437-963-66100
; Sequence 66100, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boucharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 66100
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_M74530_67084C.1
US-10-437-963-66100

```

Alignment Scores: Pred. No.: 6,75e-92 Length: 1402 Score: 821.50 Matches: 179

```

Percent Similarity: 58.85% Conservative: 87
Best Local Similarity: 39.60% Mismatches: 36
Query Match: 35.08% Indels: 36
DB: 17 Gaps: 6

US-09-857-525C-2 (1-451) x US-10-437-963-66100 (1-1402)

Qy      10  IleGluGly-CysPro-----ThrValPheGlnAlaIleCysSerLeuSe 24
Db      69  GTGGAGGAGTGGCCCATATGGGGCTCGTGGGGCCGAGTTCCAGGTGGTTCGATCTGCC 128
Qy      24  rProGlyIleHisGluTyrLysPhePheValAspGlyGluTrpArgHisAspGluArg 44
Db      129  CCGCGGGGTTTATCAGTACCGGTTTGGTGTGATGCTGCTGGCGCATGTGATGACGAA 188
Qy      44  nProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyrAs 64
Db      189  ACCTTGTGTACGGGATGAATATGGGTTGATTAGCAATGAAGTCTCTTGTGGATAATACACA 248
Qy      64  nGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspValAs 84
Db      249  TCCGGTT-----GTACA 260
Qy      84  pAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrVal----- 99
Db      261  GCCAGAACTTCCATCAGAGTAGTAGTATGGATGGAGGTACTACTATGACACAAATGCC 320
Qy      100  -SerGluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSe 119
Db      321  CCCAGATCAGTTATCTCAAAACTCAGCGCTGCAGATAGCAATTTTCCGCCATCGAGTCTC 380
Qy      119  rGluTyrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAl 139
Db      381  TGAGATATTATCATATATACCATATATGATGTTGTTCTCTGTTCTACAGAGATAGCAGT 440
Qy      139  aLeuAspIleAsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyLepr 159
Db      441  TTTGATGCTCGTCTTCTCTCAAGCAAGCCCTTTAAATCATGATGATGAGGCTTTTC 500
Qy      159  oValAlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAs 179
Db      501  TTTGGTTCTCTTTGGGATGATCAGCAGCAAAACCCGATTAACAGGCATGCTAACCGCATCCGA 560
Qy      179  pPheIleLeuLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLe 199
Db      561  TTTTGTATTAATTTTGAGAAAGCTGCAGAGAAACATTCGAACTCTTGGCCATGAGGAAT 620
Qy      199  uGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAs 219
Db      621  TGAATGCAATTCAGTGCTCTGTTGAAAGAAAGCAAAATTCAGTTTATGGGGACCTGA 680
Qy      219  pSerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAs 239
Db      681  TGTGTGCCAATTCAGAGAAAGCCATTAATTCATGTTAAGGATTCAGATATTTAAGGGA 740
Qy      239  pIleAlaValLysLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSe 259
Db      741  TGTGGCATTCGCTATTAATAGAAATGAATATCTCAGTTCCCT---ATCTTAAGCCCTC 797
Qy      259  rSerAspGlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuLysCy 279
Db      798  AAGGATTCATCAGGGATGCTTGTGCTTGTCAACCCCTTCCAGGGATTTGGAAT 857
Qy      279  sIleCysArgTyrPheLysAsnSerThrGlyAsnLeuLeuProIleLeuAsnGlnProValCy 299
Db      858  TATTGTCTCAAGCTCAAGACAGCCTGAAGGGTACTATTCTTCGAAAACCCAGATTGT 917
Qy      299  sSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAl 319
Db      918  CAGTATGCCCTATTGGTACATGTCACCATATCTGCGAAGCAGCAAGCAAGCAGCTTAC 977
Qy      319  aMetLeuArgProAsnAlaSerLeuSerAlaLeuAsnMetLeuValGlnAlaGlyVa 339

```

Db 978 AACTTCGGCGACCAAGCACTCCTCTAAATTATCATGCTGGATTGCTGTGAGATAGAGT 1037
Qy 339 lserSerIleProIleValAspAsnAspSerLeuLeuAspThrTySerArgSerAs 359
Db 1038 AAGCTCAATTCCTATAGTTGACGATAATGGCGCTCTCCTTGATGTCTACTCGCTCAGTGA 1097
Qy 359 pIleThrAlaLeuAlaLysAspLysValTyThrHisValArgLeuAspGluMetThrIle 379
Db 1098 TATCATGGCTCTTAGCAAGAATGATGTCTACACTCGTATTGAGCTTGAAACACAGGTACCGT 1157
Qy 379 eHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProp>eGlyPhePheAsnGlyG1 399
Db 1158 GGAGCATGCCITGGAGCTGCAATACACAGTG-----AATGGCGG 1196
Qy 399 nArg---CysGlnMetCysLeuLeuArgSerAspProLeuLeuLysValMetGluArgLeuAl 418
Db 1197 AAGCACTGTCTATACCTCTTGAGCACTAGTACCTCTCTGGAGGTTTGGAGCAATTGTC 1256
Qy 418 aAsnProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIle 438
Db 1257 AGCTCCAGGGGTGGCGGAGTGTCTGTTATTGAACCAAGGAGCAGATTGTGCAAGGAAT 1316
Qy 438 eIleSerLeuSerAspIlePheLysPheLeuLeu 449
Db 1317 AATCTCAATGAGGAGCGCATTTACATTCTCTCATT 1350

RESULT 14

US-10-437-963-22833
; Sequence 22833, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Audrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 22833
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27972C.1

US-10-437-963-22833

Alignment Scores:
Pred. No.: 4,06e-85 Length: 1112
Score: 766.00 Matches: 152
Percent Similarity: 95.78% Conservative: 7
Best Local Similarity: 91.57% Mismatches: 6
Query Match: 32.71% Indels: 1
DB: 17 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-437-963-22833 (1-1112)

Qy 286 AsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySer 305
Db 5 AATCTCAAGGTAATTTACCTATTATTGACCAAGCTGTATGCACAAATTCCTCTGGTACC 64
Qy 306 TrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAla 325
Db 65 TGGGTTTCAAAATAATGGTGATTCCTAATGCGCGTTCATTGGCTATGTGGCGCCCTAACACA 124
Qy 326 SerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleVal 345

Db 125 TCTCTTAGGGCTGCCTCGAATTGCTGGTTCAAGCTGGTGTGAGCTCAATACCAATTGTG 184
Qy 346 AspAspAsnAspSerLeuLeuAspThrTySerArgSerAspIleThrAlaLeuAlaLys 365
Db 185 GATGACAATGACTCACTGCTTGACACATATTCCAGAAGTGACATCACAGCTCTTGCAAAA 244
Qy 366 AspLysValTyThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeu 385
Db 245 GACAAGGTGTACACACATTCGCTAGATGAGATGACCAATTCATCAGGCTCTGCAGCTT 304
Qy 386 GlyGlnAspAlaAsnThr-PropheGlyPhePheAsnGlyGlnArgCysGlnMetCysLe 405
Db 305 GCACAGGATGCMAATTCGCCCCCTTTGGATTCTTTAAACGGGCAAAAGATGTCAGATGTGCT 364
Qy 405 nArgSerAspProLeuLeuLysValMet-SiuarGleuAlaAsnProGlyValArgArgVa 425
Db 365 CCGGTCTGACACTTTGTTGAAAGTGATGAGGAGGATTAGCTAATCTCGGGGTGGCGGTGT 424
Qy 425 lPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePh 445
Db 425 TTTCATTGTGAAGCTGTGTAGCAAAACGTGTGAGGAGGCAATAATATCACTCAGTGATATTTT 484
Qy 445 eLysPheLeuLeuSer 450
Db 485 CAAGTTCTTGCTGAGC 500

RESULT 15

US-10-425-114-8857
; Sequence 8857, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8857
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700799332_FLI

US-10-425-114-8857

Alignment Scores:
Pred. No.: 2,94e-84 Length: 1533
Score: 761.00 Matches: 173
Percent Similarity: 57.59% Conservative: 85
Best Local Similarity: 38.62% Mismatches: 152
Query Match: 32.49% Indels: 28
DB: 13 Gaps: 9

US-09-857-525C-2 (1-451) x US-10-425-114-8857 (1-1533)

Qy 13 CysPro-----ThrValPheGlnAlaIleCysSerLeuSerProGlyIle 27
Db 168 TGTCCCATGGGCTGGTGGGCGCGTTCAGGTGCTTCGATCTGCCCCCGGAGTT 227
Qy 28 HisGluTyrlsPhePheValAspGlyGluTrpArgHisAspGluArgGlnProThrIle 47
Db 228 TACCAGTACCGGTTTTTGGTTGATGGTCTCTGGAGGTGTGATGAGACAAAACCCCTTTGTA 287
Qy 48 SerGlyGluPheGlyIleValAsnThrLeuTyrlLeuThrArgGluTyrlAsnGlnIleAsn 67
Db 288 CGTGATGAATATGAGTCACTGATCAGCAATGAGTGTCTGTGGAAAAACAATGTACACCTGTT 347

QY 68 ThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspSerGluAsn 87
Db 348 GTGCGACGAGCGCTTCTATCAGAGAACT-----AATATGGAT----- 386
QY 88 PheGlnArgThrValThrLeuSerAspGlyThrVal-----SerGluGly 102
Db 387 -----AGGGTACCATTITGAAACAATAAGTCCCGGAGCCA 422
QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
Db 423 TCATCTCAGAACCCAGCATGCAATAGCAGTTATCCGCCATGTGGTCTCTGGAATATTA 482
QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 483 TTACACAATACCATATATGAGTGTTCCTCTTAGCAAGTTAACAGTTTGGACACT 542
QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaPro 162
Db 543 CAGCTTCCTCTGTTAAACAAGCATTTAAATAATGATGATGAGGGTCTTGTCTGGTTCCT 602
QY 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 603 CTTTGGGATCAGCGTCAGGAGCAACATAACAGGCATGCTCACTGCATCAGATTGTGTTTA 662
QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis 202
Db 663 ATCTTGAGAAAGTTGCAAGAAACATTCAGATTATTGGCAACGAAGAG----- 710
QY 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyr 222
Db 711 CCCATTTCCTGTTGAAAGAAAGCAAGCTACAGTTCCTATGGTGGGCGCTGATGGTGGCC 770
QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
Db 771 ATGCAGAGAGGCCATTATCCATGTTAAGGATTTCAGATTAATTTAGTGGATGGCACTG 830
QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
Db 831 ACTATAATCAGAAATGAAATATCTTCAGTTCTCT---ATCTTTAAGTGCATGGCAGATTCA 887
QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
Db 888 TCAGGGGTGCTTCTCTTAATCTTGCAACCTTACAGGGGATTTTGAAATTCCTTTGCTCG 947
QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
Db 948 AAGCTACAAGAGAGGCTGAGGGCTGTTCCTTCCTGCACAATCAGCTTCTCAGATTCTCT 1007
QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
Db 1008 ATTGGTACATGGTCTCCACATACGGGGAGGTCAAGTAGCAGGCAACTCAGAACCTTTGCTA 1067
QY 323 ProAsnAlaSerLeuSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
Db 1068 CTGAGTTCCTCTTAATACCTCCCTGGATTTCTGCTTCAAGATAGATAGCTCGATT 1127
QY 343 ProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
Db 1128 CCTATAGTTGATGACAATGGATCCCTCGGTGATGCTTACTCACTCAGTGTATCATGGCT 1187
QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
Db 1188 CTGGCAAGAAATGATGTTTATGCTCGCATCGAATCTGAACAGTGCACCGTACAAAATGCT 1247
QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArg---Cys 401
Db 1248 TTGGATGTGCATACAGGTG-----CATGCCGAGACAGCTGT 1286
QY 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly 421
Db 1287 CATACTTGTTTACAGCAGAGTACCTTGTGCGAAGTTTGGAGGGATTTGCCATTCCAGGA 1346
QY 422 ValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeu 441

Db 1347 GTGCGACGCTTGTGTTATTGAACAAAGTACCAGATTTGTGGAAGGATCTCTCATTTG 1406
QY 442 SerAspIlePheLysPheLeuLeu 449
Db 1407 AGAGACGTATTATTCATTTCCTTT 1430
Search completed: July 9, 2004, 11:24:30
Job time : 697 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2315	98.8	496	10	Q9FUY4		Q9fuy4 zea mays (m
2	2218	94.7	497	10	Q9FUY5		Q9fuy5 zea mays (m
3	2085	89.0	493	10	Q84M80		Q84m80 oryza sativ
4	1580	67.5	487	10	Q94A6		Q94a6 arabidopsis
5	1481.5	63.3	485	10	Q84P80		Q84p80 medicago tr
6	1334.5	57.0	382	10	Q9FV59		Q9fv59 arabidopsis
7	1030.5	44.0	391	10	Q04028		Q04028 arabidopsis
8	820.5	35.0	451	10	Q7XV93		Q7xv93 oryza sativ
9	405	17.3	577	5	Q86KN7		Q86kn7 dictyostell
10	376.5	16.1	443	11	Q8BIQ9		Q8biq9 mus musculu
11	374	16.0	566	11	Q91WG5		Q91wg5 mus musculu
12	369	15.8	585	5	Q9IN38		Q9in38 drosophila
13	366	15.6	906	5	Q8SX18		Q8sx18 drosophila
14	365	15.6	582	5	Q9BHL6		Q9bhl6 caenorhabdi
15	363.5	15.5	647	5	Q96613		Q96613 drosophila
16	363.5	15.5	906	5	Q9IOB2		Q9iob2 drosophila

[illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.

QY 123 NLHTCYDLPSGKVTALDINKLVKQSFHILHEQIGPVAPLWDSRFGQVGLLSPDLFIL 182
 Db :|||||
 168 SLHTCYDLLPSGKVIADINKLVKQSFHILHEQIGPVAPLWDSRFGQVGLLSPDLFIL 227
 QY 183 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 242
 Db :|||||
 228 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 287
 QY 243 KLLONGISTVPVIYSSSDGSPQLHLASLSGLKICICRYFNKSTGNLPILNQPVCSIP 302
 Db :|||||
 288 KLLONGISTVPVIYSSSDGSPQLHLASLSGLKICICRYFNKSTGNLPILNQPVCSIP 347
 QY 303 LGSWPKIGDLNRPPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLLDTYSRSDITA 362
 Db :|||||
 348 LGSWPKIGDLNRPPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLLDTYSRSDITA 407
 QY 363 LAKDKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRQCMCLRSDDLKVMERLANPGV 422
 Db :|||||
 408 LAKDKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRQCMCLRSDDLKVMERLANPGV 467
 QY 423 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 451
 Db :|||||
 468 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 496

RESULT 2

Q9FU5 PRELIMINARY; PRT; 497 AA.
 AC Q9FU5;
 DT 01-MAR-2001 (TremBLrel. 16, Created)
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE Protein kinase AKINbetagamma-1.
 CS Zea mays (Maize);
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 CX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Victoria L., Mar A., Tatjana K., Csaba K., Montserrat P.;
 RT "Domain fusion between Snf1 related kinase subunits during plant evolution."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF276085; AAG31751.1; -
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR InterPro; IPR000644; CBS_domain.
 DR Pfam; PF00571; CBS; 3.
 DR SMART; SM00116; CBS; 3.
 KW Kinase.
 SQ SEQUENCE 497 AA; 54938 MW; B9EF6A530CIAD8B CRC64;

Query Match 94.7%; Score 2218; DB 10; Length 497;
 Best Local Similarity 95.1%; Pred. No. 1.7e-179;
 Matches 427; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 3 EHLPMSPLEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTISGFGVNTLYLTRE 62
 Db :|||||
 49 EHLPMSPLEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTISGFGVNTLYLTRE 108
 QY 63 YNQINTLSSPTGSRMMMDVNDNFQRTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 122
 Db :|||||
 109 FNQINALLNPSTGSRMMMDVNDNFQRTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 168
 QY 123 NLHTCYDLLPSGKVIADINKLVKQSFHILHEQIGPVAPLWDSRFGQVGLLSPDLFIL 182
 Db :|||||
 169 NLHTCYDLLPSGKVIADINKLVKQSFHILHEQIGPVAPLWDSRFGQVGLLSPDLFIL 228
 QY 183 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 242
 Db :|||||
 229 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 288
 QY 243 KLLONGISTVPVIYSSSDGSPQLHLASLSGLKICICRYFNKSTGNLPILNQPVCSIP 302

Db :|||||
 289 KLLINDISTVPVIYSSSDGSPQLHLASLSGLKICIFRYFNKSTGNLPILNQPVCSIP 348
 QY 303 LGSWPKIGDLNRPPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLLDTYSRSDITA 362
 Db :|||||
 349 LGSWPKIGDLNRPPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLLDTYSRSDITA 408
 QY 363 LAKDKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRQCMCLRSDDLKVMERLANPGV 422
 Db :|||||
 409 LAKDKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRQCMCLRSDDLKVMERLANPGV 468
 QY 423 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 451
 Db :|||||
 469 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 497

RESULT 3

Q84M80 PRELIMINARY; PRT; 493 AA.
 AC Q84M80;
 DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Hypothetical protein OSJNBa0059G06.25.
 CN OSJNBa0059G06.25.
 CS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaeae; Oryza.
 CX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Fadrosch D.W., Tailon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNBa0059G06 genomic sequence."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell R.i.
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC096690; AAP21389.1; -
 DR InterPro; IPR000644; CBS_domain.
 DR Pfam; PF00571; CBS; 3.
 DR SMART; SM00116; CBS; 3.
 KW Hypothetical protein.
 SQ SEQUENCE 493 AA; 54493 MW; E0531CEA7481CD5F CRC64;

Query Match 89.0%; Score 2085; DB 10; Length 493;
 Best Local Similarity 87.3%; Pred. No. 3.3e-168;
 Matches 391; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

QY 3 EHLPMSPLEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTISGFGVNTLYLTRE 62
 Db :|||||
 46 EHLPMSPLEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTISGFGVNTLYLTRE 105
 QY 63 YNQINTLSSPTGSRMMMDVNDNFQRTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 122
 Db :|||||
 106 FQINTILSPSTGSRMMMDVNDNFQRTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 165
 QY 123 NLHTCYDLLPSGKVIADINKLVKQSFHILHEQIGPVAPLWDSRFGQVGLLSPDLFIL 182
 Db :|||||
 166 NLHTCYDLLPSGKVIADINKLVKQSFHILHEQIGPVAPLWDSRFGQVGLLSPDLFIL 225
 QY 183 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 242
 Db :|||||
 226 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 285

```

QY 243 KLLONGISTVPVITYSSSDGSPOLLHLSLGLKICRYFKNSTGNLPLINOPVCSIP 302
DB 286 KILONGVSTVPVIFSSSPDQSYQOLLHLSLGLKICRYFKNSTGNLPLINOPVCSIP 345
QY 303 LGSWVPKIGDNLSPRLAMLRPNASLSALNMLVQAGVSSIPVDDNDLSLDTYSRSDIT 362
DB 346 LGTWVPKIGDNLSPRLAMLRPNASLSALNMLVQAGVSSIPVDDNDLSLDTYSRSDIT 405
QY 363 LAKDKVYTHVRLDEMTTHQALQGDANTPGFPNGORCOMCLRSDDLKVMERLANPGV 422
DB 406 LAKDKVYTHVRLDEMTTHQALQGDANTPGFPNGORCOMCLRSDDLKVMERLANPGV 465
QY 423 RRVFIVEAGSKRVEGIISLSDIFKLLS 450
DB 466 RRVFIVEAGSKRVEGIISLSDIFKLLS 493

RESULT 4
Q944A6 PRELIMINARY; PRT; 487 AA.
ID Q944A6
AC Q944A6
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE At1G09020/F7G19.11.
OS Arabidopsis thaliana (Mouse-ear cress).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
CC eurosids I; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RN [2]
SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koeser E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.X., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA "Arabidopsis cDNA clones";
RL Submitted (0CT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Chen R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan J.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA "Arabidopsis ORF clones";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF439826; AAL27498.1; -.
DR EMBL; BT000625; AAN18191.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 3.
SQ SEQUENCE 487 AA; 53466 MW; 4545FE3BF2C4EBEC CRC64;

```

Query Match 67.5%; Score 1580; DB 10; Length 487;
 Best Local Similarity 67.6%; Pred. No. 2.4e-125;
 Matches 304; Conservative 64; Mismatches 74; Indels 8; Gaps 5;

```

QY 3 EHLPMSPLEGCPVTFQAIQCSLSGIIHYKFFVGEWRHDEROPTISGFGVNTLYLTRE 62
DB 45 EHVPMSPLEGCPVTFQAIQCSLSGIIHYKFFVGEWRHDEROPTISGFGVNTLYLTRE 104
QY 63 YNOINTLSSPTSGSRNMNDVNFENFQRTVTLSDGTVSEGLRVSEAAIQTISRCRVSEYL 122
DB 105 -DVNPAGFPETLG-RSNMVDV-DVFLRTADPS-----QEAVPRMSGVDVLSRHRISVLL 157

```

```

QY 123 NLHTCYDLLPDGSKVIALDINLPVKQSPHILHEGIPVAPLWDSPRGQVGLLSPLDFIL 182
DB 158 STRTAYELLPSGSKVIALDINLPVKQSPHILHEGIPVAPLWDSPRGQVGLLSPLDFIL 217
QY 183 ILRELETHGSLNLTBEQLETHITISAWKEAKQTNCRNDQWMP-QQHLVHATPYESLRDIA 241
DB 218 ILRELETHGSLNLTBEQLETHITISAWKEAKQTNCRNDQWMP-QQHLVHATPYESLRDIA 277
QY 242 VKLLONGISTVPVITYSSSDGSPOLLHLSLGLKICRYFKNSTGNLPLINOPVCSIP 301
DB 278 LKILQNVAAVPIVITYSSSDGSPOLLHLSLGLKICRYFKNSTGNLPLINOPVCSIP 337
QY 302 PLGSWVPKIGDNLSPRLAMLRPNASLSALNMLVQAGVSSIPVDDNDLSLDTYSRSDIT 361
DB 338 PLGTWVPKIGDNLSPRLAMLRPNASLSALNMLVQAGVSSIPVDDNDLSLDTYSRSDIT 397
QY 362 ALAKDKVYTHVRLDEMTTHQALQGDANTPGFPNGORCOMCLRSDDLKVMERLANPG 421
DB 398 ALAKDKVYTHVRLDEMTTHQALQGDANTPGFPNGORCOMCLRSDDLKVMERLANPG 457
QY 422 VRRVFIIVEAGSKRVEGIISLSDIFKLLS 451
DB 458 VRRVFIIVEAGSKRVEGIISLSDIFKLLS 487

RESULT 5
Q84PE0 PRELIMINARY; PRT; 485 AA.
ID Q84PE0
AC Q84PE0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AKIN beta-gamma.
OS Medicago truncatula (Barrel medic).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
CC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
CX NCBI_TaxID=3880;
RN [1]
RN [2]
SEQUENCE FROM N.A.
RA Buitink J., Thomas M., Gissot L., Leprince O.;
RA "Differential expression patterns of beta and gamma subunit isoforms
RA suggest the involvement of different SnRK1 complexes in desiccation
RA tolerance, osmotic stress and starvation in germinating seeds of
RA Medicago truncatula";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY247268; AAO61673.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 3.
DR SMART; SM00116; CBS; 3.
SQ SEQUENCE 485 AA; 53607 MW; 18E908AEBE916F48 CRC64;

```

Query Match 63.3%; Score 1481.5; DB 10; Length 485;
 Best Local Similarity 64.1%; Pred. No. 5.4e-117;
 Matches 287; Conservative 65; Mismatches 83; Indels 13; Gaps 4;

```

QY 3 EHLPMSPLEGCPVTFQAIQCSLSGIIHYKFFVGEWRHDEROPTISGFGVNTLYLTRE 62
DB 49 EHVPMSPLEGCPVTFQAIQCSLSGIIHYKFFVGEWRHDEROPTISGFGVNTLYLTRE 108
QY 63 YNOINTLSSPTSGSRNMNDVNFENFQRTVTLSDGTVSEGLRVSEAAIQTISRCRVSEYL 122
DB 109 PDILPAISAEI-SSRSHMEVDNDVFGH-----AANRMSPSDLEVSRRISKFL 158
QY 123 NLHTCYDLLPDGSKVIALDINLPVKQSPHILHEGIPVAPLWDSPRGQVGLLSPLDFIL 182
DB 159 SEHTAYDLLPSGSKVIALDINLPVKQSPHILHEGIPVAPLWDSPRGQVGLLSPLDFIL 218
QY 183 ILRELETHGSLNLTBEQLETHITISAWKEAKQTNCRNDQWMP-QQHLVHATPYESLRDIA 241
DB 219 ILKELGTHGSLNLTBEQLETHITISAWKEAKQTNCRNDQWMP-QQHLVHATPYESLRDIA 278

```



```

RESULT 8
Q7XV93
ID Q7XV93 PRELIMINARY; PRT; 451 AA.
AC Q7XV93
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSUNB0012808.3 protein.
GN OSUNB0012808.3
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.F., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Xu X.H.,
RA Li T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying X., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL606610; CAD40779.1; -
SQ SEQUENCE 451 AA; 50481 MW; 2BD7E89F568BF692 CRC64;

Query Match 35.0%; Score 820.5; DB 10; Length 451;
Best Local Similarity 39.5%; Pred. No. 6.2e-61;
Matches 177; Conservative 86; Mismatches 150; Indels 35; Gaps 6;

QY 13 CP-----TVFQAICSLSPGHEHYKFFVGGWRHDERQPTISGEFIVNTLYLTREYNQIN 67
DB 27 CPWGLGAEFQVVDLPFGVYQVREFLVGGWRCDKTPCVRDEYGLISNEVLVDNTHPV- 85

QY 68 TLSPSTPGSRMMMDVNEFNQRTVTLSDGT-----SEGLRVSEAAIQSRGVSEYL 122
DB 86 -----VQETSIRVSMDEGTLITMPPDQLSQNSGVQIAIFRERVSEIL 130

QY 123 NLHTCYDLPDSGVIALDNLVQKSFHLLHEOGIIPVAPLWDSEFGQFVGLSLPLDFIL 182
DB 131 LEMTIYDVVPSKIAVLDAFLPVKQAFKIMHDEGLSIVLWDQOQTVTGMLTASDFVL 190

QY 183 ILRELETHGSLNLTBEQLETHITSAWKAQKQTNGRDNDQWRPQOHLVHATPYESLRDIIV 242
DB 191 ILAKLQNRNIRTLGHELEHMSVSAWKAQKQFYCGPDVAAIQRRPLIHVXDSNLRDVAL 250

QY 243 KLLONGISTVPVIYSSSSDGSFPOLLHLASLSGLKICRYFNKSTGNLPLINQVCSIP 302
DB 251 AITRNEISSVP-IPKPTSSGMPLGLGLATLPGIVKFCIKLQEPBGYSFLNQIVSMP 309

QY 303 LGSWVPKIGLNSRPLAMRNASISSALNMLVOAGVSSIPVDDNDSLDITYSRSDITA 362
DB 310 IGWSPHTGKASNRQLRTSRPSTPNSCLLLEDVRSSIPVDDNGALLDVLSQIMA 369

QY 363 LAKDKVYTHVRLDMTHQALQGLQDANPPGFFNGOR-COMCLURSDPLLKWMERLANPG 421
DB 370 LGRNDVYTRIELEQVTVHEALELQYQV-----NGRRHCHTCLSTSTFLEVLQLSAPG 422

QY 422 VRRVFIKAGSKRVEGIISLSDIFKELL 449
DB 423 VRRVVVIEPRSRFVQGIISLRDAFTFLI 450

RESULT 9
Q86KN7
ID Q86KN7 PRELIMINARY; PRT; 577 AA.
AC Q86KN7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

QY 34 VDGEWRHDERQPTISGEFIVNTLYL-----TREYNQNTL-----S 70
DB 130 VDGEKGNKESQSPPGNDQILNNNNMFFKDIITSLPSTDNKSTNNNNNNENPLKQTISS 189

QY 71 SP-----TPG-----SRMMVDVNEFCETVTLSDGTVSEGLRV----- 106
DB 190 SPKSTITTTTSTTTTTPSLSSLSNNNNNNNNNNYINHSISTVSEGLENNLKGSI 249

QY 107 -----SEMAIQISRCRVSEYLNLTCTDILDPGKVIADLNLVQKSFHILHEQGIPIV 160
DB 250 KKIDSEKITEEGKQVFNELKGTCTDIPISGVVLDTKLAVKSAFALENGIKS 309

QY 161 APLWDSFRGQVGLSLPLDFI-LILRELETHGSLNLTBEQLETHITSA-WKAQKQTNGRN 218
DB 310 APLWNSRQHDFTGMTVTSDFIDILLYYKPKSNINIFQDMGHRHRETETWREI----- 361

QY 219 DSQWRPQOHLVHATPYESLRDIIVKLLONGISTVPVIYSSSSDGSFPOLLHLASLSGLK 278
DB 362 -SVERPSS-LISTPPTNLXDAASLLCYKHLRFLPVVDKDTN-----SILHILTHSRILA 415

QY 279 CICRYEKSTGNLPLINQVCSIPILGSWVPKIGLNSRPLAMRNASISSALNMLVOAG 338
DB 416 FMMKSF-----PQLEKLSIFGS-----LGITGTFATVTVMTHTPLVEVLELLSEKK 464

QY 339 VSSIPVDDNDSD-LDITYSRSDITAKADKYVTHVRLDMTHQAL-----QIGQDANTFP 393
DB 465 ISAVPIIDSETSKXIVDVYKSDVTLMKQGIILSPDLN-LPVHQLVSLTFTTKLWQ----- 517

QY 394 GFFNGQRQCMCLURSDPLLKWMERLANPGVRRVFIKAGSKRVEGIISLSDIFKELL 450
DB 518 ---RPEQITCTREFLDGLDVIERCIRKRVHRLVCIDS-SKKVEGLSLSDILNLYLLN 570

RESULT 10
Q8BIQ9
ID Q8BIQ9 PRELIMINARY; PRT; 443 AA.
AC Q8BIQ9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to AMP activated protein kinase gamma 1.
GN 2410051C13RIK.
OS Mus musculus (Mouse).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002);
 DR EMBL; AK032238; BAC27775.1;
 DR MGD; MGI:1920950; 2410051C13Rik.
 DR InterPro; IPR006644; CBS_domain.
 DR Pfam; PF00571; CBS; 4.
 DR SMART; SM00116; CBS; 4.
 SQ SEQUENCE 443 AA; 49980 MW; AC336EAD3FE35E28 CRC64;

Query Match 16.1%; Score 376.5; DB 11; Length 443;
 Best Local Similarity 28.0%; Pred. No. 2.9e-23;
 Matches 117; Conservative 79; Mismatches 143; Indels 79; Gaps 17;

QY 61 REYNQINTLSPSPSTPGSRMNDVNFQ-----RTVTL--SDGTVEG 102
 DB 62 RPESRIYASSPPDTGQRFCL-----AFQSPARPPPLASFTYHAPLRTAVLAAPGPAAG 116
 QY 103 TLRV-----SEAAIQISRCRVSEYLNLHTCYDLLPDSGKVIADINLPVKOSPHIL 153
 DB 117 MLEKLEFOEAEDESSEGVY-----RPMRSKCYDIVFTSSKLVVFTTLQVKKAFAL 170
 QY 154 HEQIPVAPLWDSFRGQVGLSLPLDLILRELETHGSNLTE-EOLETHHTISAKKAKR 212
 DB 171 VANGVRAAPLWESKKQSFVGMLTITDFINILHRY--YKSPMVQIYELEBKTIETRELVL 228
 QY 213 QTNGRNDQWRPQCHLVHATPVESLRDIAVKLLONGISTVPVIYSSSDGSPQLHLAS 272
 DB 229 Q-----ETFRP---LVNISPDASLFDAVYSLIKKHLPLVIDPISGNA-----LYILT 274
 QY 273 LSGILKICRYFKNSTGNLPILNQPVCSIPLGSWVPKIGDLSRPLMLRPNASLSALN 332
 DB 275 HKRIKFL-QLFMSDMPKPAFMKQNLDELIGTY-----HNIATFHPDTPPIKALN 324
 QY 333 MLVQAGVSSIPVDNDSLLDTYSRSDITAKAKVYTHVRLDEMTIHOALQIQDANTP 392
 DB 325 IFVERRISALPVDESCKVDIYKFDVINLAAEKTNNL---DITVTQALQHRSQ---- 377
 QY 393 FGFGNQRQCMCLRSDDLKVMERLANPGRVRFVIV-EAGSKRVEGIIISLSDIFKFL 449
 DB 378 --YFEG--VVKSKLETLETIVDRIVRAEVHRLVVVNEADS--IVGIIISLSDILQALI 429

RESULT 11
 Q91WG5 PRELIMINARY; PRT; 566 AA.
 ID Q91WG5
 AC Q91WG5
 DT 01-DEC-2001 (TrEMBLrel. 19, Created;
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN 2410051C13Rik.
 CS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Kidney;
 RA Strausberg R.;
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015283; AAH15283.1;
 DR MGD; MGI:1920950; 2410051C13Rik.
 DR InterPro; IPR006644; CBS_domain.

DR Pfam; PF00571; CBS; 4.
 DR SMART; SM00116; CBS; 4.
 RW Hypothetical protein.
 SQ SEQUENCE 566 AA; 62997 MW; 754DE4D696102C7D CRC64;

Query Match 16.0%; Score 374; DB 11; Length 566;
 Best Local Similarity 28.1%; Pred. No. 6.9e-23;
 Matches 117; Conservative 78; Mismatches 144; Indels 78; Gaps 17;

QY 61 REYNQINTLSPSPSTPGSRMNDVNFQ-----RTVTL--SDGTVEG 102
 DB 186 RPESRIYASSPPDTGQRFCL-----AFQSPARPPPLASFTYHAPLRTAVLAAPGPAAG 240
 QY 103 TLRV-----SEAAIQISRCRVSEYLNLHTCYDLLPDSGKVIADINLPVKOSPHIL 154
 DB 241 MLEKLEFOEAEDESSEGVY-----RPMRSKCYDIVFTSSKLVVFTTLQVKKAFALV 294
 QY 155 EQGIPVAPLWDSFRGQVGLSLPLDLILRELETHGSNLTE-EOLETHHTISAKKAKRQ 213
 DB 295 ANGVRAPLWESKKQSFVGMLTITDFINILHRY--YKSPMVQIYELEBKTIETRELVLQ 352
 QY 214 TNGRNDQWRPQCHLVHATPVESLRDIAVKLLONGISTVPVIYSSSDGSPQLHLASL 273
 DB 353 -----ETFRP---LVNISPDASLFDAVYSLIKKHLPLVIDPISGNA-----LYILTH 398
 QY 274 LSGILKICRYFKNSTGNLPILNQPVCSIPLGSWVPKIGDLSRPLMLRPNASLSALNM 333
 DB 399 KRILKFL-QLFMSDMPKPAFMKQNLDELIGTY-----HNIATFHPDTPPIKALNI 448
 QY 334 LVQAGVSSIPVDNDSLLDTYSRSDITAKAKVYTHVRLDEMTIHOALQIQDANTPF 393
 DB 449 FVERRISALPVDESCKVDIYKFDVINLAAEKTNNL---DITVTQALQHRSQ----- 500
 QY 394 FGFGNQRQCMCLRSDDLKVMERLANPGRVRFVIV-EAGSKRVEGIIISLSDIFKFL 449
 DB 501 -YFEG--VVKSKLETLETIVDRIVRAEVHRLVVVNEADS--IVGIIISLSDILQALI 552

RESULT 12
 Q9IN38 PRELIMINARY; PRT; 585 AA.
 ID Q9IN38
 AC Q9IN38
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE CG17299-PH.
 GN SNF4A-GAMMA OR SNF4AGAMMA OR CG5806 OR CG17299.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazek R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadiieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

367	Db	RASADLEKRRASVGAAGRGRLRGDGTLDPHRAAILFRDSRGLPVADPFLEKYNLSDEEDD	422
121	Qy	-----YLNLTCTCYDLLPDSKVIADINTLPVKQSFIHEQIGIPVAPLWDSFRGQFVGL	174
427	Db	SQIEVKFREFHKKYDIIPTSAKLAVFDQLLVKAFVALVYGVRAAPLWDSKQOQFVGM	486
175	Qy	LSPLDFILILRELETHGSNLTZEOLFTHTTSANKEAKRQTNGRNDSSQWRPQOHLVHATPY	234
487	Db	LTITDFTKIL-QMTYKSPNASMEQLEHKLDTWRSVLH-----NOVMP---LVSIGPD	535
235	Qy	ESLRDIAVKLLQNGISTVPVIYSSSSDGSPPQLHLASLSGLIKLCICRYFKNSTGMPLIL	294
536	Db	ASLYDAIKLHLSRHLRPVI-----DPATGNVLYILTHKRLILFLFLY-----	579
295	Qy	NOFVCSIFLGSWVP-----KIGDLNRPRLAMURPNASLSALNMUVQAGVSSIPVDN	348
580	Db	---INELPKPAYMOKSLRELKIGTYNN--TETADETTSLITALKKFFVERVSALPLVDS	634
349	Qy	DSGLTYSRSDITALAKDKVYTHVRLDEMITHQALQLQGDANTPFFGFNGRCQMCLRS	408
635	Db	GRLEVIYAKFDVINLAAEKTNYDL---DVSURKANEHNE-----WFEG--VQKNLDE	683
409	Qy	PLIKVMERLANPGRVRFIVEAGSKRVEGIIISDIEKFLI	449
684	Db	SLYTWERIVRAEVRHLVVWDE-NRKVIGIISLSDILLYLV	723
RESULT 14			
Q9BHL6		PRELIMINARY;	PRT; 582 AA.
ID	Q9BHL6		
AC	Q9BHL6		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
GN	Hypochemical protein Y111B2A.8.		
DB	Y111B2A.8.		
OS	Caenorhabditis elegans.		
OC	Rukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Pelodierinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RC	MEDLINE=99069613; PubMed=9851916;		
EX			
RA	none;		
RT	*Genome sequence of the nematode C.elegans: A platform for		
RL	investigating biology.*;		
RL	Science 282:2012-2018(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	Salston J.E.;		
RL	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; A0132904; CAC35836.2; -.		
KW	Hypochemical protein.		
SQ	SEQUENCE 582 AA; 64153 MW; DQFCF77F406EA2C6 CRC64;		
Query Match 15.6%; Score 365; DB 5; Length 582;			
Best Local Similarity 28.6%; Pred. No. 4.2e-22;			
Matches 104; Conservative 73; Mismatches 135; Indels 52; Gaps			
Qy	94	LSPTGVSECTLRVSEA-----ATQISRCRVSEVLMHTCYDLLPDSGKVIADINLPVKQS	149
Db	185	LSDPNDFEFTWTSRADNHDAV-----YSLFMKAHKCYDLIPTSSKLVVFDTHLPVKA	238
Qy	150	FHLHEQIGIPVAPLWDSFRGQFVGLSLPLDFILRLRELETHGSNLTE-EGLETHTISAWK	208
Db	239	FVALVYGVRAAPLWDTNQRTGMLTTTDFIKILKXHYDKGNSERIRALEDDQQISHWR	298
Qy	209	EAKRQTNGRNDSSQWRPQOHLVHATPYESLRDIANVLQNGISTVPVIYSSSSDGSFPQLL	268
Db	299	D-----QFELDGTLRP---FVYIDNESLHRAVELLCESKVHRLPLV-----DRKTGNIT	345

QY 427 IVEAGSRKVEGIIISLSDIFKELL 449
Db 443 VVDE-NEKVIIGIISLSDILLYLV 464

Search completed: July 7, 2004, 17:52:19
Job time : 48 secs

QY 269 HLASLSGILKACICRYFKNSTGNILNOPVCS---IPUGSWPKIGDLNRSPLAMLRPXA 325
Db 346 YILTHKRMFLSLYMR---DLPRFSMCTPRELIGAW---GD-----ILCCHVDY 392
QY 326 SLSSALNMLVQAGVSSIPVDDNDLSLETYSRSDITALAKDKVYTHVRLDENTHQAQL 385
Db 393 PIHDALEFLKNRVSALPLIDENGRVVDIVAKFDVISLAAESSYDKL---DCTVQEAHQH 449
QY 386 QDANTPFPGNGORCQCLRSDELLKMEELANPGVRRVPIVEAGSRKVEGIIISLSDIF 445
Db 450 RSE-----WFEG--VQTCLETESLFQVLEAIVAEVRLIVTDQ-DKKVGVVVSLSLIL 500
QY 446 KELL 449
Db 501 KNLV 504

RESULT 15
O96613 PRELIMINARY; ERT; 647 AA.
AC O96613
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SNF4/AMP-activated protein kinase gamma subunit.
GN SNF4A-GAMMA OR CG5806 OR CG17299.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoshida E.N., Benkel B.F., Fong Y., Hickey D.A.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF094764; AAC95306.1; -
DR EMBL; AF094763; AAC95305.1; -
DR FlyBase; FBgn0025803; SNF4A-gamma.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Kinase.
SQ SEQUENCE 647 AA; 71592 MW; B792BE1089730B52 CRC64;

Query Match 15.5%; Score 363.5; DB 5; Length 647;
Best Local Similarity 28.2%; Pred. No. 6.6e-22;
Matches 108; Conservative 72; Mismatches 123; Indels 74; Gaps 13;
QY 97 GTVSEGTLRVSEAAIQ-----ISGRVSE-----YLNHHTCYLLP 132
Db 126 GLRGDGLDHPHAAILFEDSRGLPVADPFLEKVNLSDEEDDSQIFVKAFRFHFKCYLLP 185
QY 133 DSGKVIALDINLPVKQSHILHEQIPVAPWDSFRGQVGLLSPDLFILIRELETHGS 192
Db 186 TSAKLWVFDQTLLVKKAFYALVYNGVRAAPLWDSKQFVGMLTITDFIKL-QMYKSP 244
QY 193 NLTEQLECHTISAWKEAKQRTNGRNDSCWEPQOHLVHATPYESLRDIKVLQNGISTV 252
Db 245 NASMEQLEEHKLDTWRRDYLH-----NQVMP---LVSIGEDASLYDAIKLIHSRIHL 294
QY 253 PVIYSSSSDGSFPOLLHLASLGILKICRYFKNSTGNILNOPVCSPIPLGSWVP--- 308
Db 295 PVI-----DPATGNVYILTHKRIILFLFLY-----INELPKPAYMKSIR 335
QY 309 --KIGDLNRSPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDLSLDTYSRSDITALAKD 366
Db 336 ELKIGTYNN--IETADETTTIIITALKKFRVRSVSLPLVDSGRLVDIYAKFDVINLAAE 393
QY 367 KYTHVRLDENTHQAQLQGDANTPGGFNGQRCQCLRSDELLKMEELANPGVRRV 426
Db 394 KTYNDL---DVSRLKANEHRNE-----WEEG--VQCNLDESLYTIMERIVRAEVHRLV 442